STIC Search Report
Biotech-Chem Library

Scientific & Technical Information Center
USPTO

STIC Database Tracking Number: 134960

TO: Maria Marvich
Location: 2b84/2c70
Art Unit: 1636
October 22, 2004

Case Serial Number: 09/762648

From: P. Sheppard
Location: Remsen Building
Phone: (571) 272-2529

sheppard@uspto.gov

Search Notes

This Page Blank (uspto)

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: October 19, 2004, 13:19:25 ; Search time 1557 Seconds
(without alignments)
303.723 Million cell updates/sec

Title: US-09-762-648-4
Perfect score: 10
Sequence: 1 UGCDGGHNMD 10

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	6.2	62.0	10	6	AR107856	AR107856 Sequence
c 2	6.2	62.0	10	6	BD238641	BD238641 Preparati
c 3	6.2	62.0	10	6	BD238894	BD238894 Preparati
4	6.2	62.0	10	6	CQ766701	CQ766701 Sequence
c 5	6.2	62.0	10	6	CQ828869	CQ828869 Sequence
6	6.2	62.0	10	6	AR234537	AR234537 Sequence
7	6.2	62.0	10	6	AR303297	AR303297 Sequence
8	6.2	62.0	11	6	A13081	A13081 Nucleotide
9	6.2	62.0	11	6	AR074496	AR074496 Sequence
10	6.2	62.0	11	6	AR081176	AR081176 Sequence
11	6.2	62.0	11	6	AR085373	AR085373 Sequence
12	6.2	62.0	11	6	AR088121	AR088121 Sequence
13	6.2	62.0	11	6	AR104280	AR104280 Sequence
c 14	6.2	62.0	11	6	AR139988	AR139988 Sequence
c 15	6.2	62.0	11	6	AR140307	AR140307 Sequence
c 16	6.2	62.0	11	6	AR140585	AR140585 Sequence
17	6.2	62.0	11	6	AR143542	AR143542 Sequence
18	6.2	62.0	11	6	AR171448	AR171448 Sequence
19	6.2	62.0	11	6	AR171619	AR171619 Sequence
20	6.2	62.0	11	6	BD243209	BD243209 MN gene a
c 21	6.2	62.0	11	6	CQ828868	CQ828868 Sequence
22	6.2	62.0	11	6	CQ833562	CQ833562 Sequence
c 23	6.2	62.0	11	6	CQ836876	CQ836876 Sequence
24	6.2	62.0	11	6	CQ837108	CQ837108 Sequence
25	6.2	62.0	11	6	CQ837113	CQ837113 Sequence
26	6.2	62.0	11	6	CQ837507	CQ837507 Sequence
c 27	6.2	62.0	11	6	AR301722	AR301722 Sequence
c 28	6.2	62.0	11	6	AX225318	AX225318 Sequence
29	6.2	62.0	11	6	AX339223	AX339223 Sequence
c 30	6.2	62.0	11	6	AX393232	AX393232 Sequence
31	6.2	62.0	11	6	AX470780	AX470780 Sequence
c 32	6.2	62.0	11	6	AX470844	AX470844 Sequence
c 33	6.2	62.0	11	6	AX471086	AX471086 Sequence
c 34	6.2	62.0	11	6	AX624499	AX624499 Sequence
c 35	6.2	62.0	11	6	AX627640	AX627640 Sequence
c 36	6.2	62.0	11	6	AX627775	AX627775 Sequence
37	6.2	62.0	11	6	AX627871	AX627871 Sequence
38	6.2	62.0	11	6	AX628442	AX628442 Sequence
c 39	6.2	62.0	11	6	AX629146	AX629146 Sequence
40	6.2	62.0	11	6	AX629397	AX629397 Sequence
41	6.2	62.0	11	6	AX630172	AX630172 Sequence
42	6.2	62.0	11	6	AX630221	AX630221 Sequence
43	6.2	62.0	11	6	AX630249	AX630249 Sequence
c 44	6.2	62.0	11	6	AX631920	AX631920 Sequence
c 45	6.2	62.0	11	6	AX632797	AX632797 Sequence

ALIGNMENTS

RESULT 1
AR107856
LOCUS AR107856 10 bp DNA linear PAT 14-FEB-2001
DEFINITION Sequence 102 from patent US 6110667.
ACCESSION AR107856
VERSION AR107856.1 GI:12823343
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 10)
AUTHORS Lopez-Nieto,C.Eduardo. and Nigam,S.Kumar.
TITLE Processes, apparatus and compositions for characterizing nucleotide sequences based on K-tuple analysis
JOURNAL Patent: US 6110667-A 102 29-AUG-2000;
FEATURES Location/Qualifiers
source 1. .10
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN

Query Match 62.0%; Score 6.2; DB 6; Length 10;
Best Local Similarity 40.0%; Pred. No. 1.4e+07;
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 1 UGCDGGHNMD 10
:||:||: ::
Db 1 TGCTGGTCAT 10

RESULT 2
BD238641/c
LOCUS BD238641 10 bp DNA linear PAT 17-JUL-2003
DEFINITION Preparation and use of superior vaccines.
ACCESSION BD238641
VERSION BD238641.1 GI:33048411
KEYWORDS JP 2002534056-A/59.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 10)

```
  AUTHORS   Roberts,B.L. and Shankara,S.
  TITLE     Preparation and use of superior vaccines
  JOURNAL   Patent: JP 2002534056-A 59 15-OCT-2002;
            GENZYME CORP
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002534056-A/59
            PD   15-OCT-2002
            PF   18-JUN-1999 JP 2000554749
            PR   19-JUN-1998 US    60/090039,19-JUN-1998 US    60/090040 PR
            19-JUN-1998 US    60/090041,19-JUN-1998 US    60/089853 PR
            19-JUN-1998 US    60/089997,19-JUN-1998 US    60/090079 PR
            19-JUN-1998 US    60/090035,19-JUN-1998 US    60/089993 PR
            19-JUN-1998 US    60/089992,19-JUN-1998 US    60/090072 PR
            19-JUN-1998 US    60/089878,19-JUN-1998 US    60/089991 PR
            19-JUN-1998 US    60/090000,19-JUN-1998 US    60/090048 PR
            19-JUN-1998 US    60/089999,19-JUN-1998 US    60/090043 PR
            19-JUN-1998 US    60/090042,19-JUN-1998 US    60/090036 PR
            19-JUN-1998 US    60/090044,19-JUN-1998 US    60/089844 PR
            19-JUN-1998 US    60/090080,19-JUN-1998 US    60/089833 PR
            19-JUN-1998 US    60/089994,19-JUN-1998 US    60/090077 PR
            19-JUN-1998 US    60/090078,19-JUN-1998 US    60/090047 PR
            19-JUN-1998 US    60/090076,19-JUN-1998 US    60/090045 PR
            08-DEC-1998 US    60/111715
            PI   BRUCE L ROBERTS,SRINIVAS SHANKARA
            PC   C12N15/09,C12N15/09,A61K39/00,A61P35/00,A61P37/04,C12N1/15, PC
              C12N1/19,
            PC   C12N1/21,C12N5/10,G01N33/15,G01N33/50,G01N33/53,G01N33/566, PC
              G01N37/00,
            PC   C12N15/00,C12N5/00,C12N15/00
            CC   Preparation and use of superior vaccines
            FH   Key             Location/Qualifiers
            FT   source          1. .10
            FT                   /organism='Homo sapiens (human)'.
FEATURES              Location/Qualifiers
     source           1. .10
                      /organism="Homo sapiens"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:9606"
ORIGIN

  Query Match             62.0%;  Score 6.2;  DB 6;  Length 10;
  Best Local Similarity   40.0%;  Pred. No. 1.4e+07;
  Matches    4;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 UGCDGGHNMD 10
              :||:||: ::
Db         10 TGCTGGCTAG 1


RESULT 3
BD238894/c
LOCUS       BD238894                  10 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Preparation and use of superior vaccines.
ACCESSION   BD238894
VERSION     BD238894.1  GI:33048664
KEYWORDS    JP 2002534056-A/312.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 10)
  AUTHORS   Roberts,B.L. and Shankara,S.
  TITLE     Preparation and use of superior vaccines
  JOURNAL   Patent: JP 2002534056-A 312 15-OCT-2002;
            GENZYME CORP
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002534056-A/312
            PD   15-OCT-2002
            PF   18-JUN-1999 JP 2000554749
            PR   19-JUN-1998 US    60/090039,19-JUN-1998 US    60/090040 PR
            19-JUN-1998 US    60/090041,19-JUN-1998 US    60/089853 PR
            19-JUN-1998 US    60/089997,19-JUN-1998 US    60/090079 PR
            19-JUN-1998 US    60/090035,19-JUN-1998 US    60/089993 PR
            19-JUN-1998 US    60/089992,19-JUN-1998 US    60/090072 PR
            19-JUN-1998 US    60/089878,19-JUN-1998 US    60/089991 PR
            19-JUN-1998 US    60/090000,19-JUN-1998 US    60/090048 PR
            19-JUN-1998 US    60/089999,19-JUN-1998 US    60/090043 PR
            19-JUN-1998 US    60/090042,19-JUN-1998 US    60/090036 PR
            19-JUN-1998 US    60/090044,19-JUN-1998 US    60/089844 PR
            19-JUN-1998 US    60/090080,19-JUN-1998 US    60/089833 PR
            19-JUN-1998 US    60/089994,19-JUN-1998 US    60/090077 PR
            19-JUN-1998 US    60/090078,19-JUN-1998 US    60/090047 PR
            19-JUN-1998 US    60/090076,19-JUN-1998 US    60/090045 PR
            08-DEC-1998 US    60/111715
            PI   BRUCE L ROBERTS,SRINIVAS SHANKARA
            PC   C12N15/09,C12N15/09,A61K39/00,A61P35/00,A61P37/04,C12N1/15, PC
              C12N1/19,
            PC   C12N1/21,C12N5/10,G01N33/15,G01N33/50,G01N33/53,G01N33/566, PC
              G01N37/00,
            PC   C12N15/00,C12N5/00,C12N15/00
            CC   Preparation and use of superior vaccines
            FH   Key             Location/Qualifiers
            FT   source          1. .10
            FT                   /organism='Homo sapiens (human)'.
FEATURES              Location/Qualifiers
     source           1. .10
                      /organism="Homo sapiens"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:9606"
ORIGIN

  Query Match             62.0%;  Score 6.2;  DB 6;  Length 10;
  Best Local Similarity   40.0%;  Pred. No. 1.4e+07;
  Matches    4;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 UGCDGGHNMD 10
              :||:||: ::
Db         10 TGCAGGAAAA 1


RESULT 4
CQ766701
LOCUS       CQ766701                  10 bp    DNA     linear   PAT 03-MAR-2004
DEFINITION  Sequence 57 from Patent WO2004005541.
ACCESSION   CQ766701
VERSION     CQ766701.1  GI:44908931
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   van Broeckhoven,C., de Jonghe,P., Timmerman,V. and Verhoeven,K.
  TITLE     Diagnostic tests for the detection of peripheral neuropathy
  JOURNAL   Patent: WO 2004005541-A 57 15-JAN-2004;
            Vlaams Interuniversitair Instituut voor Biotechnologie vz; w. (BE)
FEATURES              Location/Qualifiers
     source           1. .10
                      /organism="synthetic construct"
                      /mol_type="unassigned DNA"
                      /db_xref="taxon:32630"
                      /note="3-intron/exon, exon 2, gene RAB7"
ORIGIN

  Query Match             62.0%;  Score 6.2;  DB 6;  Length 10;
  Best Local Similarity   40.0%;  Pred. No. 1.4e+07;
  Matches    4;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 UGCDGGHNMD 10
              :||:||: ::
Db          1 TGCAGGTAAG 10


RESULT 5
CQ828869/c
LOCUS       CQ828869                  10 bp    DNA     linear   PAT 05-JUL-2004
DEFINITION  Sequence 587 from Patent WO2004053120.
```

```
ACCESSION   CQ828869
VERSION     CQ828869.1  GI:49732352
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Weihe,E., Bieller,A. and Schaefer,M.K.
  TITLE     Regulatory elements in the 5' region of the vr1 gene
  JOURNAL   Patent: WO 2004053120-A 587 24-JUN-2004;
            Gruenenthal GmbH (DE)
FEATURES             Location/Qualifiers
     source          1. .10
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="V$MYOD Q6"
ORIGIN

  Query Match             62.0%;  Score 6.2;  DB 6;  Length 10;
  Best Local Similarity   40.0%;  Pred. No. 1.4e+07;
  Matches    4;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 UGCDGGHNMD 10
              :||:||: ::
Db         10 TGCAGGTGAA 1


RESULT 6
AR234537
LOCUS       AR234537                  10 bp    DNA     linear   PAT 20-DEC-2002
DEFINITION  Sequence 71 from patent US 6458584.
ACCESSION   AR234537
VERSION     AR234537.1  GI:27277241
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 10)
  AUTHORS   Mirzabekov,A., Guschin,D.Y., Chik,V., Drobyshev,A., Fotin,A.,
            Yershov,G. and Lysov,Y.
  TITLE     Customized oligonucleotide microchips that convert multiple genetic
            information to simple patterns, are portable and reusable
  JOURNAL   Patent: US 6458584-A 71 01-OCT-2002;
FEATURES             Location/Qualifiers
     source          1. .10
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match             62.0%;  Score 6.2;  DB 6;  Length 10;
  Best Local Similarity   40.0%;  Pred. No. 1.4e+07;
  Matches    4;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 UGCDGGHNMD 10
              .:||:||: ::
Db          1 TGCGGGTCAA 10


RESULT 7
AR303297
LOCUS       AR303297                  10 bp    DNA     linear   PAT 12-JUN-2003
DEFINITION  Sequence 22 from patent US 6544736.
ACCESSION   AR303297
VERSION     AR303297.1  GI:31692073
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 10)
  AUTHORS   Shimamoto,A., Furuichi,Y., Shibata,Y., Funaki,H., Ohara,E. and
            Watahiki,M.
  TITLE     Method for synthesizing cDNA from mRNA sample
  JOURNAL   Patent: US 6544736-A 22 08-APR-2003;
FEATURES             Location/Qualifiers
     source          1. .10
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match             62.0%;  Score 6.2;  DB 6;  Length 10;
  Best Local Similarity   40.0%;  Pred. No. 1.4e+07;
  Matches    4;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 UGCDGGHNMD 10
              :||:||: ::
Db          1 TGCTGGAAAG 10


RESULT 8
A13081
LOCUS       A13081                    11 bp    DNA     linear   PAT 08-FEB-1994
DEFINITION  Nucleotide sequence 36 from patent number EP0353188.
ACCESSION   A13081
VERSION     A13081.1  GI:489599
KEYWORDS    .
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 11)
  AUTHORS   Heim,J., Meyhack,B. and Visser,J.
  TITLE     Novel expression system
  JOURNAL   Patent: EP 0353188-A 36 31-JAN-1990;
            CIBA-GEIGY AG
FEATURES             Location/Qualifiers
     source          1. .11
                     /organism="unidentified"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32644"
ORIGIN

  Query Match             62.0%;  Score 6.2;  DB 6;  Length 11;
  Best Local Similarity   40.0%;  Pred. No. 1.3e+07;
  Matches    4;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 UGCDGGHNMD 10
              :||:||: ::
Db          2 TGCAGGTCAA 11


RESULT 9
AR074496
LOCUS       AR074496                  11 bp    DNA     linear   PAT 28-AUG-2000
DEFINITION  Sequence 75 from patent US 5955075.
ACCESSION   AR074496
VERSION     AR074496.1  GI:10001251
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 11)
  AUTHORS   Zavada,J., Pastorekova,S. and Pastorek,J.
  TITLE     Method of inhibiting tumor growth using antibodies to MN protein
  JOURNAL   Patent: US 5955075-A 75 21-SEP-1999;
FEATURES             Location/Qualifiers
     source          1. .11
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match             62.0%;  Score 6.2;  DB 6;  Length 11;
  Best Local Similarity   40.0%;  Pred. No. 1.3e+07;
  Matches    4;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 UGCDGGHNMD 10
```

```
              :||:||: ::
Db          1 TGCTGGTGAG 10


RESULT 10
AR081176
LOCUS       AR081176                  11 bp    DNA     linear   PAT 31-AUG-2000
DEFINITION  Sequence 75 from patent US 5972353.
ACCESSION   AR081176
VERSION     AR081176.1  GI:10007904
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 11)
  AUTHORS   Zavada,J., Pastorekova,S. and Pastorek,J.
  TITLE     MN proteins, polypeptides, fusion proteins and fusion polypeptides
  JOURNAL   Patent: US 5972353-A 75 26-OCT-1999;
FEATURES             Location/Qualifiers
     source          1. .11
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match             62.0%;  Score 6.2;  DB 6;  Length 11;
  Best Local Similarity   40.0%;  Pred. No. 1.3e+07;
  Matches    4;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 UGCDGGHNMD 10
              :||:||: ::
Db          1 TGCTGGTGAG 10


RESULT 11
AR085373
LOCUS       AR085373                  11 bp    DNA     linear   PAT 01-SEP-2000
DEFINITION  Sequence 75 from patent US 5981711.
ACCESSION   AR085373
VERSION     AR085373.1  GI:10012142
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 11)
  AUTHORS   Zavada,J., Pastorekova,S. and Pastorek,J.
  TITLE     MN-specific antibodies and hybridomas
  JOURNAL   Patent: US 5981711-A 75 09-NOV-1999;
FEATURES             Location/Qualifiers
     source          1. .11
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match             62.0%;  Score 6.2;  DB 6;  Length 11;
  Best Local Similarity   40.0%;  Pred. No. 1.3e+07;
  Matches    4;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 UGCDGGHNMD 10
              :||:||: ::
Db          1 TGCTGGTGAG 10


RESULT 12
AR088121
LOCUS       AR088121                  11 bp    DNA     linear   PAT 07-SEP-2000
DEFINITION  Sequence 75 from patent US 5989838.
ACCESSION   AR088121
VERSION     AR088121.1  GI:10014884
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 11)
  AUTHORS   Zavada,J., Pastorekova,S. and Pastorek,J.
  TITLE     Immunological methods of detecting MN proteins and MN polypeptides
  JOURNAL   Patent: US 5989838-A 75 23-NOV-1999;
FEATURES             Location/Qualifiers
     source          1. .11
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match             62.0%;  Score 6.2;  DB 6;  Length 11;
  Best Local Similarity   40.0%;  Pred. No. 1.3e+07;
  Matches    4;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 UGCDGGHNMD 10
              :||:||: ::
Db          1 TGCTGGTGAG 10


RESULT 13
AR104280
LOCUS       AR104280                  11 bp    DNA     linear   PAT 14-FEB-2001
DEFINITION  Sequence 75 from patent US 6093548.
ACCESSION   AR104280
VERSION     AR104280.1  GI:12816988
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 11)
  AUTHORS   Zavada,J., Pastorekova,S. and Pastorek,J.
  TITLE     Detection and quantitation of MN-specific antibodies
  JOURNAL   Patent: US 6093548-A 75 25-JUL-2000;
FEATURES             Location/Qualifiers
     source          1. .11
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match             62.0%;  Score 6.2;  DB 6;  Length 11;
  Best Local Similarity   40.0%;  Pred. No. 1.3e+07;
  Matches    4;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 UGCDGGHNMD 10
              :||:||: ::
Db          1 TGCTGGTGAG 10


RESULT 14
AR139988/c
LOCUS       AR139988                  11 bp    DNA     linear   PAT 16-JUN-2001
DEFINITION  Sequence 99 from patent US 6207417.
ACCESSION   AR139988
VERSION     AR139988.1  GI:14482484
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 11)
  AUTHORS   Zsebo,K.M., Bosselman,R.A., Suggs,S.V. and Martin,F.H.
  TITLE     DNA encoding stem cell factor
  JOURNAL   Patent: US 6207417-A 99 27-MAR-2001;
FEATURES             Location/Qualifiers
     source          1. .11
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match             62.0%;  Score 6.2;  DB 6;  Length 11;
  Best Local Similarity   40.0%;  Pred. No. 1.3e+07;
  Matches    4;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 UGCDGGHNMD 10
              :||:||: ::
```

```
Db         11 TGCAGGAGAT 2


RESULT 15
AR140307/c
LOCUS       AR140307                 11 bp    DNA     linear   PAT 16-JUN-2001
DEFINITION  Sequence 99 from patent US 6207454.
ACCESSION   AR140307
VERSION     AR140307.1  GI:14482803
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 11)
  AUTHORS   Zsebo,K.M., Bosselman,R.A., Suggs,S.V. and Martin,F.H.
  TITLE     Method for enhancing the effciency of gene transfer with stem cell
            factor (SCF) polypeptide
  JOURNAL   Patent: US 6207454-A 99 27-MAR-2001;
FEATURES             Location/Qualifiers
     source          1. .11
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match             62.0%;  Score 6.2;  DB 6;  Length 11;
  Best Local Similarity   40.0%;  Pred. No. 1.3e+07;
  Matches    4;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 UGCDGGHNMD 10
              :||:||: ::
Db         11 TGCAGGAGAT 2


Search completed: October 19, 2004, 15:31:01
Job time : 1560 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: October 19, 2004, 11:15:35 ; Search time 290 Seconds
(without alignments)
181.015 Million cell updates/sec

Title: US-09-762-648-4
Perfect score: 10
Sequence: 1 UGCDGGHNMD 10

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_23Sep04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	6.2	62.0	10	2	AAQ96908	Aaq96908 HIV-1 NL4
c	2	6.2	62.0	10	2	AAQ96474	Aaq96474 HIV-1 NL4
	3	6.2	62.0	10	2	AAT29372	Aat29372 5'-primer
	4	6.2	62.0	10	2	AAV45971	Aav45971 Biosensor
	5	6.2	62.0	10	2	AAX32311	Aax32311 Radioacti
	6	6.2	62.0	10	2	AAX32304	Aax32304 Transcrip
	7	6.2	62.0	10	2	AAX32307	Aax32307 Radioacti
	8	6.2	62.0	10	2	AAX32305	Aax32305 Radioacti
	9	6.2	62.0	10	2	AAX32309	Aax32309 Radioacti
	10	6.2	62.0	10	2	AAX32310	Aax32310 Radioacti
	11	6.2	62.0	10	2	AAX32308	Aax32308 Radioacti
c	12	6.2	62.0	10	3	AAZ77884	Aaz77884 Human den
c	13	6.2	62.0	10	3	AAZ77631	Aaz77631 Human den
	14	6.2	62.0	10	3	AAZ81050	Aaz81050 Metastati
	15	6.2	62.0	10	3	AAZ85761	Aaz85761 Metastati
c	16	6.2	62.0	10	3	AAZ79910	Aaz79910 Human den
	17	6.2	62.0	10	3	AAA95172	Aaa95172 Primer #1
	18	6.2	62.0	10	3	AAC58840	Aac58840 Human HLA
	19	6.2	62.0	10	5	AAF36075	Aaf36075 Yeast NOR
	20	6.2	62.0	10	6	AAS99948	Aas99948 Even-skip
	21	6.2	62.0	10	6	AAD42664	Aad42664 Human BUB
c	22	6.2	62.0	10	6	ABK92648	Abk92648 Primer-ex
	23	6.2	62.0	10	6	ABL52050	Abl52050 Human SLC
c	24	6.2	62.0	10	6	AAS16764	Aas16764 Human APO
	25	6.2	62.0	10	6	ABK96151	Abk96151 Human CYP
	26	6.2	62.0	10	6	ADH22211	Adh22211 Primer ex
	27	6.2	62.0	10	10	ADD71266	Add71266 Mouse ET
	28	6.2	62.0	10	10	ADD71292	Add71292 Human ET
	29	6.2	62.0	10	10	ACA63200	Aca63200 Human ALD
	30	6.2	62.0	10	10	ABX15890	Abx15890 Human BUB
c	31	6.2	62.0	10	12	ADN89079	Adn89079 Hyperlipi
c	32	6.2	62.0	10	12	ADQ30294	Adq30294 Human VR1
c	33	6.2	62.0	11	2	AAZ18993	Aaz18993 Murine MR
c	34	6.2	62.0	11	3	AAA13719	Aaa13719 Rat stem
	35	6.2	62.0	11	3	AAA16597	Aaa16597 Human MN
	36	6.2	62.0	11	3	AAA52516	Aaa52516 Human MN
	37	6.2	62.0	11	3	AAA92135	Aaa92135 Human Lhx
c	38	6.2	62.0	11	4	AAH41359	Aah41359 Recombina
c	39	6.2	62.0	11	4	AAS04138	Aas04138 Synthetic
c	40	6.2	62.0	11	4	AAF89118	Aaf89118 Stem cell
c	41	6.2	62.0	11	5	AAH23917	Aah23917 Synthetic
c	42	6.2	62.0	11	5	AAS04240	Aas04240 Synthetic
c	43	6.2	62.0	11	5	AAS10475	Aas10475 Synthetic
	44	6.2	62.0	11	6	ABQ86602	Abq86602 Human ski
c	45	6.2	62.0	11	6	ABQ86666	Abq86666 Human ski

ALIGNMENTS

RESULT 1
AAQ96908/c
ID AAQ96908 standard; DNA; 10 BP.
XX
AC AAQ96908;
XX
DT 16-OCT-2003 (revised)
DT 26-MAR-1996 (first entry)
XX
DE HIV-1 NL4-3 nef gene nucleotide deletion 503.
XX
KW HIV-1; AIDS; attenuation; vaccine; nef gene; avirulence; ss.
XX
OS Human immunodeficiency virus 1.
XX
PN WO9521912-A1.
XX
PD 17-AUG-1995.
XX
PF 14-FEB-1995; 95WO-AU000063.
XX
PR 14-FEB-1994; 94AU-00003864.
PR 21-FEB-1994; 94AU-00004002.
PR 23-DEC-1994; 94AU-00000284.
XX
PA (MACF-) MACFARLANE BURNET CENT MEDICAL.
PA (AURE-) AUSTRALIAN RED CROSS SOC.
XX
PI Deacon NJ, Learmont JC, Mcphee DA, Crowe S, Cooper D;
XX
DR WPI; 1995-293115/38.
XX
PT New non-pathogenic HIV-1 strain carrying a deletion in its nef gene or
PT LTR region - can be used in a vaccine to inhibit/reduce productive
PT infection in an individual by a pathogenic strain.
XX
PS Claim 13; Page 194; 301pp; English.
XX
CC Attenuation of pathogenic HIV-1 strain NL4-3 involves deletion of 1 or
CC more decanucleotides (AAQ96406-Q97018) from the nef gene and/or 1 or more
CC decanucleotides (AAQ97019-Q97166) from the LTR region; the sequence of
CC AAQ96406 corresponds to nucleotides 1-10 of the nef gene (AAQ96141). The
CC resulting avirulent HIV strains are still capable of inducing an immune
CC response in humans, and enable the generation of therapeutic, diagnostic

```
CC   and targeting agents against HIV-1 infection. (Updated on 16-OCT-2003 to
CC   standardise OS field)
XX
SQ   Sequence 10 BP; 2 A; 3 C; 3 G; 2 T; 0 U; 0 Other;

  Query Match             62.0%;  Score 6.2;  DB 2;  Length 10;
  Best Local Similarity   40.0%;  Pred. No. 6.7e+05;
  Matches    4;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 UGCDGGHNMD 10
              :||:||: ::
Db         10 TGCAGGCTCA 1


RESULT 2
AAQ96474/c
ID   AAQ96474 standard; DNA; 10 BP.
XX
AC   AAQ96474;
XX
DT   16-OCT-2003  (revised)
DT   20-MAR-1996  (first entry)
XX
DE   HIV-1 NL4-3 nef gene nucleotide deletion 69.
XX
KW   HIV-1; AIDS; attenuation; vaccine; nef gene; avirulence; ss.
XX
OS   Human immunodeficiency virus 1.
XX
PN   WO9521912-A1.
XX
PD   17-AUG-1995.
XX
PF   14-FEB-1995;   95WO-AU000063.
XX
PR   14-FEB-1994;   94AU-00003864.
PR   21-FEB-1994;   94AU-00004002.
PR   23-DEC-1994;   94AU-00000284.
XX
PA   (MACF-) MACFARLANE BURNET CENT MEDICAL.
PA   (AURE-) AUSTRALIAN RED CROSS SOC.
XX
PI   Deacon NJ,  Learmont JC,  Mcphee DA,  Crowe S,  Cooper D;
XX
DR   WPI; 1995-293115/38.
XX
PT   New non-pathogenic HIV-1 strain carrying a deletion in its nef gene or
PT   LTR region - can be used in a vaccine to inhibit/reduce productive
PT   infection in an individual by a pathogenic strain.
XX
PS   Claim 13; Page 189; 301pp; English.
XX
CC   Attenuation of pathogenic HIV-1 strain NL4-3 involves deletion of 1 or
CC   more decanucleotides (AAQ96406-Q97018) from the nef gene and/or 1 or more
CC   decanucleotides (AAQ97019-Q97166) from the LTR region; the sequence of
CC   AAQ96406 corresponds to nucleotides 1-10 of the nef gene (AAQ96141). The
CC   resulting avirulent HIV strains are still capable of inducing an immune
CC   response in humans, and enable the generation of therapeutic, diagnostic
CC   and targeting agents against HIV-1 infection. (Updated on 16-OCT-2003 to
CC   standardise OS field)
XX
SQ   Sequence 10 BP; 3 A; 3 C; 3 G; 1 T; 0 U; 0 Other;

  Query Match             62.0%;  Score 6.2;  DB 2;  Length 10;
  Best Local Similarity   40.0%;  Pred. No. 6.7e+05;
  Matches    4;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 UGCDGGHNMD 10
              :||:||: ::
Db         10 TGCTGGCTCA 1


RESULT 3
AAT29372
ID   AAT29372 standard; DNA; 10 BP.
XX
AC   AAT29372;
XX
DT   25-MAR-2003  (revised)
DT   28-JUN-1996  (first entry)
XX
DE   5'-primer for mammalian G-protein coupled receptor coding sequences.
XX
KW   5'-primer; mammalian; G-protein coupled receptor; PCR primer kit;
KW   characterisation; biological samples; PCR amplification; indexing;
KW   identification; cloning; analysis; genes; genome mapping;
KW   disease diagnosis; ss.
XX
OS   Synthetic.
XX
PN   WO9531574-A1.
XX
PD   23-NOV-1995.
XX
PF   12-MAY-1995;   95WO-US006032.
XX
PR   16-MAY-1994;   94US-00242887.
XX
PA   (BGHM ) BRIGHAM & WOMENS HOSPITAL.
XX
PI   Lopeznieto CE,  Nigam SK;
XX
DR   WPI; 1996-010958/01.
XX
PT   Characterisation of nucleotide sequences using primer pairs - by PCR
PT   amplification and indexing of amplification prods. w.r.t. primers used
PT   for genome mapping and disease diagnosis.
XX
PS   Claim 46; Page 55; 72pp; English.
XX
CC   The 5'-primers AAT29262-382, and the complementary 3'-primers derived
CC   from them, which target mammalian G-protein coupled receptor coding
CC   sequences, together comprise a PCR primer kit. The kit is used in a new
CC   method for the characterisation of nucleic acid sequences obtd. from
CC   mammalian biological samples, which comprises PCR amplification and
CC   indexing of the prods. w.r.t the primer pair that hybridised to its
CC   delineating subsequences. The method may be used in the identification,
CC   cloning and analysis of genes, e.g. in genome mapping, and disease
CC   diagnosis. (Updated on 25-MAR-2003 to correct PI field.)
XX
SQ   Sequence 10 BP; 1 A; 2 C; 3 G; 4 T; 0 U; 0 Other;

  Query Match             62.0%;  Score 6.2;  DB 2;  Length 10;
  Best Local Similarity   40.0%;  Pred. No. 6.7e+05;
  Matches    4;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 UGCDGGHNMD 10
              :||:||: ::
Db          1 TGCTGGTCAT 10


RESULT 4
AAV45971
ID   AAV45971 standard; DNA; 10 BP.
XX
AC   AAV45971;
XX
DT   08-OCT-1998  (first entry)
XX
DE   Biosensor oligonucleotide #71.
XX
KW   Biosensor; microchip; detection; identification; hybridisation; ss.
XX
OS   Synthetic.
XX
PN   WO9828444-A2.
```

XX
PD 02-JUL-1998.
XX
PF 19-DEC-1997; 97WO-US023778.
XX
PR 23-DEC-1996; 96US-00780026.
XX
PA (UYCH-) UNIV CHICAGO.
XX
PI Mirzabekov A, Guschin DY, Shik V, Fotin A, Yershov G, Lysov Y;
XX
DR WPI; 1998-377672/32.
XX
PT New customised oligonucleotide microchips - comprising a matrix of
PT oligonucleotides for hybridisation for identifying nucleic acid sequences
PT in samples.
XX
PS Claim 22; Fig 9; 76pp; English.
XX
CC AAV45901-V45974 are oligonucleotides which are used as biosensors in a
CC method for using a microchip to identify a sequence in a nucleic acid of
CC a sample. Such chips can be used to identifying nucleotide sequences in
CC samples such as air, water, soil, blood, cells, tissue, tissue culture
CC and food. The same microchip can be used for hybridisation for more than
CC 20-30 times, without any noticeable deterioration of the hybridisation
CC signal. Also, parallel hybridisation of nucleic acids in a sample to many
CC oligonucleotides on a microchip is possible, allowing replication and
CC standardisation
XX
SQ Sequence 10 BP; 2 A; 2 C; 4 G; 2 T; 0 U; 0 Other;

Query Match 62.0%; Score 6.2; DB 2; Length 10;
Best Local Similarity 40.0%; Pred. No. 6.7e+05;
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 1 UGCDGGHNMD 10
:||:||: ::
Db 1 TGCGGGTCAA 10

RESULT 5
AAX32311
ID AAX32311 standard; RNA; 10 BP.
XX
AC AAX32311;
XX
DT 16-JUN-1999 (first entry)
XX
DE Radioactivator isolate no: 7.
XX
KW Transcriptional regulator; DNA-binding moiety; promoter; transcription;
KW initiation; elongation; gene expression; cancer; anemia; erythrocyte;
KW riboactivator; R10 library; ss.
XX
OS Synthetic.
XX
PN WO9910487-A2.
XX
PD 04-MAR-1999.
XX
PF 26-AUG-1998; 98WO-US017691.
XX
PR 27-AUG-1997; 97US-0056857P.
XX
PA (HARD) HARVARD COLLEGE.
PA (UYBO-) UNIV BOSTON.
XX
PI Jarrell KA, Saha S, Ptashne M;
XX
DR WPI; 1999-204663/17.
XX
PT New RNA transcriptional regulators - used for modulating gene expression
PT in vitro or in vivo, e.g. as therapeutic agents for treating e.g. cancer
PT or anaemia.
XX
PS Example 2; Page 23; 57pp; English.
XX
CC The invention provides novel transcriptional regulators (TR) that are
CC comprised of RNA molecules. The TR comprises a DNA-binding moiety and an
CC RNA linked to the DNA binding moiety, where the RNA has TR activity.
CC Methods of identifying such RNA TRs are also provided. The RNA TRs alter
CC the rate and/or the extent of transcription from a promoter when they are
CC delivered to a site that is operationally linked to that promoter. The
CC TRs can affect transcription initiation, elongation, reinitiation,
CC termination and pausing. They can be used in e.g. bacterial cells, yeast
CC cells, mammalian cells, insect cells, plant cells, reptile cells,
CC celenorate cells, and protozoan cells. They can be used as agents for
CC controlling gene expression, e.g. to modulate gene expression in vivo in
CC order to alleviate or correct a disease state, e.g. cancer, anemia and
CC other disorders related to erythrocyte production. Sequences AAX32305-312
CC represent radioactivator isolate sequences obtained from a R10 library
XX
SQ Sequence 10 BP; 2 A; 2 C; 3 G; 0 T; 3 U; 0 Other;

Query Match 62.0%; Score 6.2; DB 2; Length 10;
Best Local Similarity 50.0%; Pred. No. 6.7e+05;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 UGCDGGHNMD 10
|||:||: ::
Db 1 UGCUGGAUCA 10

RESULT 6
AAX32304
ID AAX32304 standard; RNA; 10 BP.
XX
AC AAX32304;
XX
DT 16-JUN-1999 (first entry)
XX
DE Transcriptional regulatory RNA sequence.
XX
KW Transcriptional regulator; DNA-binding moiety; promoter; transcription;
KW initiation; elongation; gene expression; cancer; anemia; erythrocyte; ss.
XX
OS Synthetic.
XX
PN WO9910487-A2.
XX
PD 04-MAR-1999.
XX
PF 26-AUG-1998; 98WO-US017691.
XX
PR 27-AUG-1997; 97US-0056857P.
XX
PA (HARD) HARVARD COLLEGE.
PA (UYBO-) UNIV BOSTON.
XX
PI Jarrell KA, Saha S, Ptashne M;
XX
DR WPI; 1999-204663/17.
XX
PT New RNA transcriptional regulators - used for modulating gene expression
PT in vitro or in vivo, e.g. as therapeutic agents for treating e.g. cancer
PT or anaemia.
XX
PS Claim 9; Page 45; 57pp; English.
XX
CC The invention provides novel transcriptional regulators (TR) that are
CC comprised of RNA molecules. The TR comprises a DNA-binding moiety and an
CC RNA linked to the DNA binding moiety, where the RNA has TR activity.
CC Methods of identifying such RNA TRs are also provided. The RNA TRs alter
CC the rate and/or the extent of transcription from a promoter when they are
CC delivered to a site that is operationally linked to that promoter. The
CC TRs can affect transcription initiation, elongation, reinitiation,

```
CC   termination and pausing. They can be used in e.g. bacterial cells, yeast
CC   cells, mammalian cells, insect cells, plant cells, reptile cells,
CC   celenorate cells, and protozoan cells. They can be used as agents for
CC   controlling gene expression, e.g. to modulate gene expression in vivo in
CC   order to alleviate or correct a disease state, e.g. cancer, anemia and
CC   other disorders related to erythrocyte production. The present sequence
CC   represents a transcriptional regulatory RNA sequence
XX
SQ   Sequence 10 BP; 0 A; 1 C; 3 G; 0 T; 1 U; 5 Other;

  Query Match             62.0%;  Score 6.2;  DB 2;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 6.7e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 UGCDGGHNMD 10
              ||||||||||
Db          1 UGCDGGHNMD 10


RESULT 7
AAX32307
ID   AAX32307 standard; RNA; 10 BP.
XX
AC   AAX32307;
XX
DT   16-JUN-1999  (first entry)
XX
DE   Radioactivator isolate no: 3.
XX
KW   Transcriptional regulator; DNA-binding moiety; promoter; transcription;
KW   initiation; elongation; gene expression; cancer; anemia; erythrocyte;
KW   riboactivator; R10 library; ss.
XX
OS   Synthetic.
XX
PN   WO9910487-A2.
XX
PD   04-MAR-1999.
XX
PF   26-AUG-1998;   98WO-US017691.
XX
PR   27-AUG-1997;   97US-0056857P.
XX
PA   (HARD ) HARVARD COLLEGE.
PA   (UYBO-) UNIV BOSTON.
XX
PI   Jarrell KA,  Saha S,  Ptashne M;
XX
DR   WPI; 1999-204663/17.
XX
PT   New RNA transcriptional regulators - used for modulating gene expression
PT   in vitro or in vivo, e.g. as therapeutic agents for treating e.g. cancer
PT   or anaemia.
XX
PS   Example 2; Page 23; 57pp; English.
XX
CC   The invention provides novel transcriptional regulators (TR) that are
CC   comprised of RNA molecules. The TR comprises a DNA-binding moiety and an
CC   RNA linked to the DNA binding moiety, where the RNA has TR activity.
CC   Methods of identifying such RNA TRs are also provided. The RNA TRs alter
CC   the rate and/or the extent of transcription from a promoter when they are
CC   delivered to a site that is operationally linked to that promoter. The
CC   TRs can affect transcription initiation, elongation, reinitiation,
CC   termination and pausing. They can be used in e.g. bacterial cells, yeast
CC   cells, mammalian cells, insect cells, plant cells, reptile cells,
CC   celenorate cells, and protozoan cells. They can be used as agents for
CC   controlling gene expression, e.g. to modulate gene expression in vivo in
CC   order to alleviate or correct a disease state, e.g. cancer, anemia and
CC   other disorders related to erythrocyte production. Sequences AAX32305-312
CC   represent radioactivator isolate sequences obtained from a R10 library
XX
SQ   Sequence 10 BP; 1 A; 2 C; 4 G; 0 T; 3 U; 0 Other;

  Query Match             62.0%;  Score 6.2;  DB 2;  Length 10;
  Best Local Similarity   50.0%;  Pred. No. 6.7e+05;
  Matches    5;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 UGCDGGHNMD 10
              |||:||: ::
Db          1 UGCGGGUCAU 10


RESULT 8
AAX32305
ID   AAX32305 standard; RNA; 10 BP.
XX
AC   AAX32305;
XX
DT   16-JUN-1999  (first entry)
XX
DE   Radioactivator isolate no: 1.
XX
KW   Transcriptional regulator; DNA-binding moiety; promoter; transcription;
KW   initiation; elongation; gene expression; cancer; anemia; erythrocyte;
KW   riboactivator; R10 library; ss.
XX
OS   Synthetic.
XX
PN   WO9910487-A2.
XX
PD   04-MAR-1999.
XX
PF   26-AUG-1998;   98WO-US017691.
XX
PR   27-AUG-1997;   97US-0056857P.
XX
PA   (HARD ) HARVARD COLLEGE.
PA   (UYBO-) UNIV BOSTON.
XX
PI   Jarrell KA,  Saha S,  Ptashne M;
XX
DR   WPI; 1999-204663/17.
XX
PT   New RNA transcriptional regulators - used for modulating gene expression
PT   in vitro or in vivo, e.g. as therapeutic agents for treating e.g. cancer
PT   or anaemia.
XX
PS   Example 2; Page 23; 57pp; English.
XX
CC   The invention provides novel transcriptional regulators (TR) that are
CC   comprised of RNA molecules. The TR comprises a DNA-binding moiety and an
CC   RNA linked to the DNA binding moiety, where the RNA has TR activity.
CC   Methods of identifying such RNA TRs are also provided. The RNA TRs alter
CC   the rate and/or the extent of transcription from a promoter when they are
CC   delivered to a site that is operationally linked to that promoter. The
CC   TRs can affect transcription initiation, elongation, reinitiation,
CC   termination and pausing. They can be used in e.g. bacterial cells, yeast
CC   cells, mammalian cells, insect cells, plant cells, reptile cells,
CC   celenorate cells, and protozoan cells. They can be used as agents for
CC   controlling gene expression, e.g. to modulate gene expression in vivo in
CC   order to alleviate or correct a disease state, e.g. cancer, anemia and
CC   other disorders related to erythrocyte production. Sequences AAX32305-312
CC   represent radioactivator isolate sequences obtained from a R10 library
XX
SQ   Sequence 10 BP; 1 A; 2 C; 5 G; 0 T; 2 U; 0 Other;

  Query Match             62.0%;  Score 6.2;  DB 2;  Length 10;
  Best Local Similarity   50.0%;  Pred. No. 6.7e+05;
  Matches    5;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 UGCDGGHNMD 10
              |||:||: ::
Db          1 UGCGGGUACG 10


RESULT 9
```

```
AAX32309
ID   AAX32309 standard; RNA; 10 BP.
XX
AC   AAX32309;
XX
DT   16-JUN-1999  (first entry)
XX
DE   Radioactivator isolate no: 5.
XX
KW   Transcriptional regulator; DNA-binding moiety; promoter; transcription;
KW   initiation; elongation; gene expression; cancer; anemia; erythrocyte;
KW   riboactivator; R10 library; ss.
XX
OS   Synthetic.
XX
PN   WO9910487-A2.
XX
PD   04-MAR-1999.
XX
PF   26-AUG-1998;   98WO-US017691.
XX
PR   27-AUG-1997;   97US-0056857P.
XX
PA   (HARD ) HARVARD COLLEGE.
PA   (UYBO-) UNIV BOSTON.
XX
PI   Jarrell KA,  Saha S,  Ptashne M;
XX
DR   WPI; 1999-204663/17.
XX
PT   New RNA transcriptional regulators - used for modulating gene expression
PT   in vitro or in vivo, e.g. as therapeutic agents for treating e.g. cancer
PT   or anaemia.
XX
PS   Example 2; Page 23; 57pp; English.
XX
CC   The invention provides novel transcriptional regulators (TR) that are
CC   comprised of RNA molecules. The TR comprises a DNA-binding moiety and an
CC   RNA linked to the DNA binding moiety, where the RNA has TR activity.
CC   Methods of identifying such RNA TRs are also provided. The RNA TRs alter
CC   the rate and/or the extent of transcription from a promoter when they are
CC   delivered to a site that is operationally linked to that promoter. The
CC   TRs can affect transcription initiation, elongation, reinitiation,
CC   termination and pausing. They can be used in e.g. bacterial cells, yeast
CC   cells, mammalian cells, insect cells, plant cells, reptile cells,
CC   celenorate cells, and protozoan cells. They can be used as agents for
CC   controlling gene expression, e.g. to modulate gene expression in vivo in
CC   order to alleviate or correct a disease state, e.g. cancer, anemia and
CC   other disorders related to erythrocyte production. Sequences AAX32305-312
CC   represent radioactivator isolate sequences obtained from a R10 library
XX
SQ   Sequence 10 BP; 2 A; 2 C; 4 G; 0 T; 2 U; 0 Other;

  Query Match             62.0%;  Score 6.2;  DB 2;  Length 10;
  Best Local Similarity   50.0%;  Pred. No. 6.7e+05;
  Matches    5;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 UGCDGGHNMD 10
              |||:||: ::
Db          1 UGCGGGAUCA 10


RESULT 10
AAX32310
ID   AAX32310 standard; RNA; 10 BP.
XX
AC   AAX32310;
XX
DT   16-JUN-1999  (first entry)
XX
DE   Radioactivator isolate no: 6.
XX
KW   Transcriptional regulator; DNA-binding moiety; promoter; transcription;
KW   initiation; elongation; gene expression; cancer; anemia; erythrocyte;
KW   riboactivator; R10 library; ss.
XX
OS   Synthetic.
XX
PN   WO9910487-A2.
XX
PD   04-MAR-1999.
XX
PF   26-AUG-1998;   98WO-US017691.
XX
PR   27-AUG-1997;   97US-0056857P.
XX
PA   (HARD ) HARVARD COLLEGE.
PA   (UYBO-) UNIV BOSTON.
XX
PI   Jarrell KA,  Saha S,  Ptashne M;
XX
DR   WPI; 1999-204663/17.
XX
PT   New RNA transcriptional regulators - used for modulating gene expression
PT   in vitro or in vivo, e.g. as therapeutic agents for treating e.g. cancer
PT   or anaemia.
XX
PS   Example 2; Page 23; 57pp; English.
XX
CC   The invention provides novel transcriptional regulators (TR) that are
CC   comprised of RNA molecules. The TR comprises a DNA-binding moiety and an
CC   RNA linked to the DNA binding moiety, where the RNA has TR activity.
CC   Methods of identifying such RNA TRs are also provided. The RNA TRs alter
CC   the rate and/or the extent of transcription from a promoter when they are
CC   delivered to a site that is operationally linked to that promoter. The
CC   TRs can affect transcription initiation, elongation, reinitiation,
CC   termination and pausing. They can be used in e.g. bacterial cells, yeast
CC   cells, mammalian cells, insect cells, plant cells, reptile cells,
CC   celenorate cells, and protozoan cells. They can be used as agents for
CC   controlling gene expression, e.g. to modulate gene expression in vivo in
CC   order to alleviate or correct a disease state, e.g. cancer, anemia and
CC   other disorders related to erythrocyte production. Sequences AAX32305-312
CC   represent radioactivator isolate sequences obtained from a R10 library
XX
SQ   Sequence 10 BP; 1 A; 2 C; 4 G; 0 T; 3 U; 0 Other;

  Query Match             62.0%;  Score 6.2;  DB 2;  Length 10;
  Best Local Similarity   50.0%;  Pred. No. 6.7e+05;
  Matches    5;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 UGCDGGHNMD 10
              |||:||: ::
Db          1 UGCAGGUUCG 10


RESULT 11
AAX32308
ID   AAX32308 standard; RNA; 10 BP.
XX
AC   AAX32308;
XX
DT   16-JUN-1999  (first entry)
XX
DE   Radioactivator isolate no: 4.
XX
KW   Transcriptional regulator; DNA-binding moiety; promoter; transcription;
KW   initiation; elongation; gene expression; cancer; anemia; erythrocyte;
KW   riboactivator; R10 library; ss.
XX
OS   Synthetic.
XX
PN   WO9910487-A2.
XX
PD   04-MAR-1999.
XX
PF   26-AUG-1998;   98WO-US017691.
```

XX
PR 27-AUG-1997; 97US-0056857P.
XX
PA (HARD) HARVARD COLLEGE.
PA (UYBO-) UNIV BOSTON.
XX
PI Jarrell KA, Saha S, Ptashne M;
XX
DR WPI; 1999-204663/17.
XX
PT New RNA transcriptional regulators - used for modulating gene expression
PT in vitro or in vivo, e.g. as therapeutic agents for treating e.g. cancer
PT or anaemia.
XX
PS Example 2; Page 23; 57pp; English.
XX
CC The invention provides novel transcriptional regulators (TR) that are
CC comprised of RNA molecules. The TR comprises a DNA-binding moiety and an
CC RNA linked to the DNA binding moiety, where the RNA has TR activity.
CC Methods of identifying such RNA TRs are also provided. The RNA TRs alter
CC the rate and/or the extent of transcription from a promoter when they are
CC delivered to a site that is operationally linked to that promoter. The
CC TRs can affect transcription initiation, elongation, reinitiation,
CC termination and pausing. They can be used in e.g. bacterial cells, yeast
CC cells, mammalian cells, insect cells, plant cells, reptile cells,
CC celenorate cells, and protozoan cells. They can be used as agents for
CC controlling gene expression, e.g. to modulate gene expression in vivo in
CC order to alleviate or correct a disease state, e.g. cancer, anemia and
CC other disorders related to erythrocyte production. Sequences AAX32305-312
CC represent radioactivator isolate sequences obtained from a R10 library
XX
SQ Sequence 10 BP; 0 A; 2 C; 5 G; 0 T; 3 U; 0 Other;

Query Match 62.0%; Score 6.2; DB 2; Length 10;
Best Local Similarity 50.0%; Pred. No. 6.7e+05;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 UGCDGGHNMD 10
||||:||: ::
Db 1 UGCGGGUUCG 10

RESULT 12
AAZ77884/c
ID AAZ77884 standard; DNA; 10 BP.
XX
AC AAZ77884;
XX
DT 10-APR-2000 (first entry)
XX
DE Human dendritic cell SAGE tag, SEQ ID NO:312.
XX
KW SAGE tag; serial analysis of gene expression; antigen-presenting cell;
KW APC; monocyte-derived dendritic cell; differential gene expression;
KW immunostimulatory cofactor; costimulatory factor; CTL;
KW cytotoxic T-lymphocyte; tumour antigen; immunotherapy; anticancer; ss.
XX
OS Homo sapiens.
XX
PN WO9965924-A2.
XX
PD 23-DEC-1999.
XX
PF 18-JUN-1999; 99WO-US013800.
XX
PR 19-JUN-1998; 98US-0089833P.
PR 19-JUN-1998; 98US-0089844P.
PR 19-JUN-1998; 98US-0089853P.
PR 19-JUN-1998; 98US-0089878P.
PR 19-JUN-1998; 98US-0089991P.
PR 19-JUN-1998; 98US-0089992P.
PR 19-JUN-1998; 98US-0089993P.
PR 19-JUN-1998; 98US-0089994P.
PR 19-JUN-1998; 98US-0089997P.
PR 19-JUN-1998; 98US-0089999P.
PR 19-JUN-1998; 98US-0090000P.
PR 19-JUN-1998; 98US-0090035P.
PR 19-JUN-1998; 98US-0090036P.
PR 19-JUN-1998; 98US-0090039P.
PR 19-JUN-1998; 98US-0090040P.
PR 19-JUN-1998; 98US-0090041P.
PR 19-JUN-1998; 98US-0090042P.
PR 19-JUN-1998; 98US-0090043P.
PR 19-JUN-1998; 98US-0090044P.
PR 19-JUN-1998; 98US-0090045P.
PR 19-JUN-1998; 98US-0090047P.
PR 19-JUN-1998; 98US-0090048P.
PR 19-JUN-1998; 98US-0090072P.
PR 19-JUN-1998; 98US-0090076P.
PR 19-JUN-1998; 98US-0090077P.
PR 19-JUN-1998; 98US-0090078P.
PR 19-JUN-1998; 98US-0090079P.
PR 19-JUN-1998; 98US-0090080P.
PR 08-DEC-1998; 98US-0111715P.
XX
PA (GENZ) GENZYME CORP.
PA (ROBE/) ROBERTS B L.
PA (SHAN/) SHANKARA S.
XX
PI Roberts BL, Shankara S;
XX
DR WPI; 2000-106077/09.
XX
PT Isolated polynucleotides differentially expressed in antigen-presenting
PT cells, useful in gene vaccines against cancer.
XX
PS Claim 1; Page 73; 130pp; English.
XX
CC Sequences AAZ77573-Z79709 represent SAGE (serial analysis of gene
CC expression) tags used to identify mRNA transcripts encoding
CC immunostimulatory cofactor proteins which are preferentially or
CC differentially expressed in monocyte-derived dendritic cells compared
CC with monocytes. Some of the transcripts correspond to known genes or ESTs
CC (expressed sequence tags) which were previously unknown to be
CC preferentially or differentially expressed in dendritic cells, while
CC other transcripts correspond to novel genes. Antigen-presenting cell
CC (APC)-associated costimulatory factors play an important role in the
CC activation of the cytotoxic immune response, particularly against tumour
CC cells. Tumour antigen presentation via the MHC (major histocompatibility
CC complex) and subsequent recognition by T-cell receptors is alone
CC insufficient to activate a robust cytotoxic immune response that can lyse
CC the tumour cells, immunostimulatory cofactors also being required for
CC efficient activation of cytotoxic T-lymphocytes (CTLs). Nucleic acid
CC sequences identified using the SAGE tags have several potential uses.
CC They may be used in vaccines to induce an immune response, particularly
CC against a tumour antigen; to modulate the genotype of an APC; to screen
CC for agents that modulate expression of differentially expressed genes in
CC an APC; and as hybridisation probes/amplification primers for the
CC diagnosis, prognosis and monitoring of diseases related to abnormal
CC expression of these genes. Detection of the dendritic cell differentially
CC expressed genes, or of their encoded proteins, can be used to identify
CC cells as belonging to the monocyte lineage. Cells containing these genes
CC can be used in active immunotherapy (or to stimulate production of a
CC population of antigen-specific effector cells) and vectors containing
CC them are used in gene therapy. Co-administration of tumour antigens and
CC APC-associated costimulatory factors ensures adequate antigen
CC presentation to endogenous APCs and upregulates the APCs for the
CC presentation of co-stimulatory signals, migration to T cell-rich sites,
CC secretion of T cell growth factors and secretion of chemokines for
CC recruitment of immune effector cells
XX
SQ Sequence 10 BP; 1 A; 3 C; 1 G; 5 T; 0 U; 0 Other;

Query Match 62.0%; Score 6.2; DB 3; Length 10;
Best Local Similarity 40.0%; Pred. No. 6.7e+05;
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 UGCDGGHNMD 10
              :||:||: ::
Db         10 TGCAGGAAAA 1
```

RESULT 13
AAZ77631/c
ID AAZ77631 standard; DNA; 10 BP.
XX
AC AAZ77631;
XX
DT 10-APR-2000 (first entry)
XX
DE Human dendritic cell SAGE tag, SEQ ID NO:59.
XX
KW SAGE tag; serial analysis of gene expression; antigen-presenting cell;
KW APC; monocyte-derived dendritic cell; differential gene expression;
KW immunostimulatory cofactor; costimulatory factor; CTL;
KW cytotoxic T-lymphocyte; tumour antigen; immunotherapy; anticancer; ss.
XX
OS Homo sapiens.
XX
PN WO9965924-A2.
XX
PD 23-DEC-1999.
XX
PF 18-JUN-1999; 99WO-US013800.
XX
PR 19-JUN-1998; 98US-0089833P.
PR 19-JUN-1998; 98US-0089844P.
PR 19-JUN-1998; 98US-0089853P.
PR 19-JUN-1998; 98US-0089878P.
PR 19-JUN-1998; 98US-0089991P.
PR 19-JUN-1998; 98US-0089992P.
PR 19-JUN-1998; 98US-0089993P.
PR 19-JUN-1998; 98US-0089994P.
PR 19-JUN-1998; 98US-0089997P.
PR 19-JUN-1998; 98US-0089999P.
PR 19-JUN-1998; 98US-0090000P.
PR 19-JUN-1998; 98US-0090035P.
PR 19-JUN-1998; 98US-0090036P.
PR 19-JUN-1998; 98US-0090039P.
PR 19-JUN-1998; 98US-0090040P.
PR 19-JUN-1998; 98US-0090041P.
PR 19-JUN-1998; 98US-0090042P.
PR 19-JUN-1998; 98US-0090043P.
PR 19-JUN-1998; 98US-0090044P.
PR 19-JUN-1998; 98US-0090045P.
PR 19-JUN-1998; 98US-0090047P.
PR 19-JUN-1998; 98US-0090048P.
PR 19-JUN-1998; 98US-0090072P.
PR 19-JUN-1998; 98US-0090076P.
PR 19-JUN-1998; 98US-0090077P.
PR 19-JUN-1998; 98US-0090078P.
PR 19-JUN-1998; 98US-0090079P.
PR 19-JUN-1998; 98US-0090080P.
PR 08-DEC-1998; 98US-0111715P.
XX
PA (GENZ) GENZYME CORP.
PA (ROBE/) ROBERTS B L.
PA (SHAN/) SHANKARA S.
XX
PI Roberts BL, Shankara S;
XX
DR WPI; 2000-106077/09.
XX
PT Isolated polynucleotides differentially expressed in antigen-presenting
PT cells, useful in gene vaccines against cancer.
XX
PS Claim 1; Page 65; 130pp; English.
XX
CC Sequences AAZ77573-Z79709 represent SAGE (serial analysis of gene
CC expression) tags used to identify mRNA transcripts encoding
CC immunostimulatory cofactor proteins which are preferentially or
CC differentially expressed in monocyte-derived dendritic cells compared
CC with monocytes. Some of the transcripts correspond to known genes or ESTs
CC (expressed sequence tags) which were previously unknown to be
CC preferentially or differentially expressed in dendritic cells, while
CC other transcripts correspond to novel genes. Antigen-presenting cell
CC (APC)-associated costimulatory factors play an important role in the
CC activation of the cytotoxic immune response, particularly against tumour
CC cells. Tumour antigen presentation via the MHC (major histocompatibility
CC complex) and subsequent recognition by T-cell receptors is alone
CC insufficient to activate a robust cytotoxic immune response that can lyse
CC the tumour cells, immunostimulatory cofactors also being required for
CC efficient activation of cytotoxic T-lymphocytes (CTLs). Nucleic acid
CC sequences identified using the SAGE tags have several potential uses.
CC They may be used in vaccines to induce an immune response, particularly
CC against a tumour antigen; to modulate the genotype of an APC; to screen
CC for agents that modulate expression of differentially expressed genes in
CC an APC; and as hybridisation probes/amplification primers for the
CC diagnosis, prognosis and monitoring of diseases related to abnormal
CC expression of these genes. Detection of the dendritic cell differentially
CC expressed genes, or of their encoded proteins, can be used to identify
CC cells as belonging to the monocyte lineage. Cells containing these genes
CC can be used in active immunotherapy (or to stimulate production of a
CC population of antigen-specific effector cells) and vectors containing
CC them are used in gene therapy. Co-administration of tumour antigens and
CC APC-associated costimulatory factors ensures adequate antigen
CC presentation to endogenous APCs and upregulates the APCs for the
CC presentation of co-stimulatory signals, migration to T cell-rich sites,
CC secretion of T cell growth factors and secretion of chemokines for
CC recruitment of immune effector cells
XX
SQ Sequence 10 BP; 3 A; 4 C; 2 G; 1 T; 0 U; 0 Other;

  Query Match   62.0%; Score 6.2; DB 3; Length 10;
  Best Local Similarity   40.0%; Pred. No. 6.7e+05;
  Matches   4; Conservative   5; Mismatches   1; Indels   0; Gaps   0;

```
Qy          1 UGCDGGHNMD 10
              :||:||: ::
Db         10 TGCTGGCTAG 1
```

RESULT 14
AAZ81050
ID AAZ81050 standard; DNA; 10 BP.
XX
AC AAZ81050;
XX
DT 07-APR-2000 (first entry)
XX
DE Metastatic breast tumour cell upregulated transcript tag #284.
XX
KW Human; metastatic breast tumour tissue; breast cancer; tag; primer;
KW non-metastatic breast tumour tissue; gene therapy; anticancer;
KW antimetastatic; vaccine; diagnosis; ss.
XX
OS Homo sapiens.
XX
PN WO9965928-A2.
XX
PD 23-DEC-1999.
XX
PF 18-JUN-1999; 99WO-US013647.
XX
PR 19-JUN-1998; 98US-0089853P.
PR 19-JUN-1998; 98US-0089997P.
PR 19-JUN-1998; 98US-0090039P.
PR 19-JUN-1998; 98US-0090040P.
PR 19-JUN-1998; 98US-0090041P.
XX
PA (GENZ) GENZYME CORP.
PA (ROBE/) ROBERTS B L.

PA (SHAN/) SHANKARA S.
XX
PI Roberts BL, Shankara S;
XX
DR WPI; 2000-106079/09.
XX
PT Isolated polynucleotides differentially expressed between metastatic and
PT non-metastatic breast cancer cells, useful for diagnosis, prevention and
PT treatment of cancer.
XX
PS Claim 1; Page 65; 219pp; English.
XX
CC AAZ80767 to AAZ83941 represent tags corresponding to distinct transcripts
CC that are preferentially transcribed in the metastatic breast tumour
CC tissue (i.e. are upregulated in metastatic breast tumour cells). AAZ83942
CC to AAZ86677 represent tags corresponding to distinct transcripts that are
CC preferentially transcribed in the primary or non-metastatic breast tumour
CC tissue (i.e. are downregulated in metastatic breast tumour cells). These
CC transcripts can be used for diagnosis, prognosis, monitoring and
CC treatment of breast cancer, particularly where metastatic. Diagnosis is
CC by standard immunoassays or hybridisation/amplification reactions.
CC Compounds that modulate expression of the transcripts are potentially
CC useful for treatment of (metastatic) breast cancer, while promoters from
CC the transcripts are used to direct expression, in selected cell types, of
CC e.g. therapeutic genes (also ribozymes or antisense sequences),
CC particularly an antigen-encoding sequence for use in gene or cell-based
CC vaccines. Polypeptides encoded by the transcripts are also useful in
CC vaccines; for diagnosing breast cancer and for raising specific
CC antibodies (Ab). Ab are used to detect the polypeptides or as therapeutic
CC agents. Host cells that produce the polypeptides can be used to expand
CC and isolate populations of educated, antigen-specific immune effecter
CC cells, e.g. cytotoxic T lymphocytes, and these used for adoptive
CC immunotherapy
XX
SQ Sequence 10 BP; 2 A; 2 C; 3 G; 3 T; 0 U; 0 Other;

Query Match 62.0%; Score 6.2; DB 3; Length 10;
Best Local Similarity 40.0%; Pred. No. 6.7e+05;
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 1 UGCDGGHNMD 10
:||:||: ::
Db 1 TGCAGGTACT 10

RESULT 15
AAZ85761
ID AAZ85761 standard; DNA; 10 BP.
XX
AC AAZ85761;
XX
DT 07-APR-2000 (first entry)
XX
DE Metastatic breast tumour cell downregulated transcript tag #4995.
XX
KW Human; metastatic breast tumour tissue; breast cancer; tag; primer;
KW non-metastatic breast tumour tissue; gene therapy; anticancer;
KW antimetastatic; vaccine; diagnosis; ss.
XX
OS Homo sapiens.
XX
PN WO9965928-A2.
XX
PD 23-DEC-1999.
XX
PF 18-JUN-1999; 99WO-US013647.
XX
PR 19-JUN-1998; 98US-0089853P.
PR 19-JUN-1998; 98US-0089997P.
PR 19-JUN-1998; 98US-0090039P.
PR 19-JUN-1998; 98US-0090040P.
PR 19-JUN-1998; 98US-0090041P.
XX
PA (GENZ) GENZYME CORP.
PA (ROBE/) ROBERTS B L.
PA (SHAN/) SHANKARA S.
XX
PI Roberts BL, Shankara S;
XX
DR WPI; 2000-106079/09.
XX
PT Isolated polynucleotides differentially expressed between metastatic and
PT non-metastatic breast cancer cells, useful for diagnosis, prevention and
PT treatment of cancer.
XX
PS Claim 1; Page 192; 219pp; English.
XX
CC AAZ80767 to AAZ83941 represent tags corresponding to distinct transcripts
CC that are preferentially transcribed in the metastatic breast tumour
CC tissue (i.e. are upregulated in metastatic breast tumour cells). AAZ83942
CC to AAZ86677 represent tags corresponding to distinct transcripts that are
CC preferentially transcribed in the primary or non-metastatic breast tumour
CC tissue (i.e. are downregulated in metastatic breast tumour cells). These
CC transcripts can be used for diagnosis, prognosis, monitoring and
CC treatment of breast cancer, particularly where metastatic. Diagnosis is
CC by standard immunoassays or hybridisation/amplification reactions.
CC Compounds that modulate expression of the transcripts are potentially
CC useful for treatment of (metastatic) breast cancer, while promoters from
CC the transcripts are used to direct expression, in selected cell types, of
CC e.g. therapeutic genes (also ribozymes or antisense sequences),
CC particularly an antigen-encoding sequence for use in gene or cell-based
CC vaccines. Polypeptides encoded by the transcripts are also useful in
CC vaccines; for diagnosing breast cancer and for raising specific
CC antibodies (Ab). Ab are used to detect the polypeptides or as therapeutic
CC agents. Host cells that produce the polypeptides can be used to expand
CC and isolate populations of educated, antigen-specific immune effecter
CC cells, e.g. cytotoxic T lymphocytes, and these used for adoptive
CC immunotherapy
XX
SQ Sequence 10 BP; 3 A; 1 C; 4 G; 2 T; 0 U; 0 Other;

Query Match 62.0%; Score 6.2; DB 3; Length 10;
Best Local Similarity 40.0%; Pred. No. 6.7e+05;
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 1 UGCDGGHNMD 10
:||:||: ::
Db 1 TGCTGGAGAA 10

Search completed: October 19, 2004, 15:04:54
Job time : 293 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: October 19, 2004, 14:46:35 ; Search time 57 Seconds
(without alignments)
124.700 Million cell updates/sec

Title: US-09-762-648-4
Perfect score: 10
Sequence: 1 UGCDGGHNMD 10

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
c 1	6.2	62.0	10	3	US-08-388-353-70	Sequence 70, Appl
c 2	6.2	62.0	10	3	US-08-388-353-504	Sequence 504, App
c 3	6.2	62.0	10	3	US-08-488-551B-70	Sequence 70, Appl
c 4	6.2	62.0	10	3	US-08-488-551B-504	Sequence 504, App
c 5	6.2	62.0	10	3	US-08-488-551B-822	Sequence 822, App
6	6.2	62.0	10	3	US-08-522-384-102	Sequence 102, App
7	6.2	62.0	10	4	US-09-261-115-71	Sequence 71, Appl
8	6.2	62.0	10	4	US-09-508-753B-22	Sequence 22, Appl
9	6.2	62.0	11	2	US-08-481-658B-75	Sequence 75, Appl
10	6.2	62.0	11	2	US-08-477-504A-75	Sequence 75, Appl
11	6.2	62.0	11	2	US-08-486-756A-75	Sequence 75, Appl
12	6.2	62.0	11	2	US-08-485-862B-75	Sequence 75, Appl
13	6.2	62.0	11	3	US-08-787-739-75	Sequence 75, Appl
14	6.2	62.0	11	3	US-08-487-077A-75	Sequence 75, Appl
15	6.2	62.0	11	3	US-08-485-863A-75	Sequence 75, Appl
16	6.2	62.0	11	3	US-08-485-049D-75	Sequence 75, Appl
c 17	6.2	62.0	11	3	US-08-482-918-99	Sequence 99, Appl
c 18	6.2	62.0	11	3	US-09-224-681-99	Sequence 99, Appl
c 19	6.2	62.0	11	3	US-08-336-728A-99	Sequence 99, Appl
20	6.2	62.0	11	3	US-09-178-115-75	Sequence 75, Appl
21	6.2	62.0	11	3	US-09-177-776-75	Sequence 75, Appl
c 22	6.2	62.0	11	4	US-09-249-155A-303	Sequence 303, App
c 23	6.2	62.0	11	4	US-09-635-251-99	Sequence 99, Appl
24	6.2	62.0	11	4	US-09-772-719B-75	Sequence 75, Appl
25	6.2	62.0	12	1	US-08-035-928-18	Sequence 18, Appl
26	6.2	62.0	12	2	US-08-441-887A-177	Sequence 177, App
c 27	6.2	62.0	12	2	US-08-993-118-6	Sequence 6, Appli
c 28	6.2	62.0	12	3	US-08-845-528C-6	Sequence 6, Appli
c 29	6.2	62.0	12	3	US-09-620-926-1	Sequence 1, Appli
30	6.2	62.0	12	4	US-09-528-404-7	Sequence 7, Appli
c 31	6.2	62.0	12	4	US-09-066-281B-6	Sequence 6, Appli
c 32	6.2	62.0	12	4	US-09-574-117A-9	Sequence 9, Appli
c 33	6.2	62.0	12	4	US-09-468-433C-6	Sequence 6, Appli
c 34	6.2	62.0	13	6	5196516-5	Patent No. 5196516
c 35	6.2	62.0	14	1	US-08-332-747-9	Sequence 9, Appli
c 36	6.2	62.0	14	1	US-08-332-747-17	Sequence 17, Appl
c 37	6.2	62.0	14	3	US-08-998-099-339	Sequence 339, App
38	6.2	62.0	14	3	US-08-998-099-368	Sequence 368, App
39	6.2	62.0	14	4	US-09-475-947A-278	Sequence 278, App
c 40	6.2	62.0	15	1	US-08-319-492B-413	Sequence 413, App
c 41	6.2	62.0	15	1	US-08-363-240A-36	Sequence 36, Appl
c 42	6.2	62.0	15	1	US-08-363-240A-69	Sequence 69, Appl
c 43	6.2	62.0	15	1	US-08-363-240A-543	Sequence 543, App
c 44	6.2	62.0	15	1	US-08-363-240A-592	Sequence 592, App
c 45	6.2	62.0	15	1	US-08-363-240A-751	Sequence 751, App

ALIGNMENTS

RESULT 1
US-08-388-353-70/c
; Sequence 70, Application US/08388353
; Patent No. 6010895
; GENERAL INFORMATION:
; APPLICANT: Deacon, Nicholas J.
; APPLICANT: Learmont, Jennifer C.
; APPLICANT: McPhee, Dale A.
; APPLICANT: Crowe, Suzanne
; APPLICANT: Cooper, David
; TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
; NUMBER OF SEQUENCES: 800
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/388,353
; FILING DATE: 14-FEB-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: DiGiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 9606
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 70:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-388-353-70

Query Match 62.0%; Score 6.2; DB 3; Length 10;
Best Local Similarity 40.0%; Pred. No. 8.9e+04;
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 1 UGCDGGHNMD 10

```
                :||:||: ::
Db          10 TGCTGGCTCA 1


RESULT 2
US-08-388-353-504/c
; Sequence 504, Application US/08388353
; Patent No. 6010895
;  GENERAL INFORMATION:
;    APPLICANT:  Deacon, Nicholas J.
;    APPLICANT:  Learmont, Jennifer C.
;    APPLICANT:  McPhee, Dale A.
;    APPLICANT:  Crowe, Suzanne
;    APPLICANT:  Cooper, David
;    TITLE OF INVENTION:  NON-PATHOGENIC STRAINS OF HIV-1
;    NUMBER OF SEQUENCES:  800
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Scully, Scott, Murphy & Presser
;      STREET:  400 Garden City Plaza
;      CITY:  Garden City
;      STATE:  New York
;      COUNTRY:  United States
;      ZIP:  11530
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/388,353
;      FILING DATE:  14-FEB-1995
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DiGiglio, Frank S.
;      REGISTRATION NUMBER:  31,346
;      REFERENCE/DOCKET NUMBER:  9606
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (516) 742-4343
;      TELEFAX:  (516) 742-4366
;      TELEX:  230 901 SANS UR
;  INFORMATION FOR SEQ ID NO:  504:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-388-353-504

  Query Match             62.0%;  Score 6.2;  DB 3;  Length 10;
  Best Local Similarity   40.0%;  Pred. No. 8.9e+04;
  Matches    4;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 UGCDGGHNMD 10
              :||:||: ::
Db          10 TGCAGGCTCA 1


RESULT 3
US-08-488-551B-70/c
; Sequence 70, Application US/08488551B
; Patent No. 6015661
;  GENERAL INFORMATION:
;    APPLICANT:  Nicholas J. Deacon
;    APPLICANT:  Dale A. McPhee
;    APPLICANT:  David Cooper
;    TITLE OF INVENTION:  NON-PATHOGENIC STRAINS OF HIV-1
;    NUMBER OF SEQUENCES:  841
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SCULLY, SCOTT, MURPHY & PRESSER
;      STREET:  400 GARDEN CITY PLAZA
;      CITY:  GARDEN CITY
;      STATE:  NEW YORK
;      COUNTRY:  U.S.A.
;      ZIP:  11530-0299
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/488,551B
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PM3864 (AU)
;      FILING DATE:  14-FEB-1994
;      APPLICATION NUMBER:  PM4002 (AU)
;      FILING DATE:  21-FEB-1994
;      APPLICATION NUMBER:  PN0284 (AU)
;      FILING DATE:  23-DEC-1994
;      APPLICATION NUMBER:  US 08/388,353
;      FILING DATE:  14-FEB-1995
;      APPLICATION NUMBER:  PN3021/95
;      FILING DATE:  17-MAY-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  FRANK S. DIGIGLIO
;      REFERENCE/DOCKET NUMBER:  9606Z
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (516) 742-4343
;      TELEFAX:  (516) 742-4366
;  INFORMATION FOR SEQ ID NO:  70:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
US-08-488-551B-70

  Query Match             62.0%;  Score 6.2;  DB 3;  Length 10;
  Best Local Similarity   40.0%;  Pred. No. 8.9e+04;
  Matches    4;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 UGCDGGHNMD 10
              :||:||: ::
Db          10 TGCTGGCTCA 1


RESULT 4
US-08-488-551B-504/c
; Sequence 504, Application US/08488551B
; Patent No. 6015661
;  GENERAL INFORMATION:
;    APPLICANT:  Nicholas J. Deacon
;    APPLICANT:  Dale A. McPhee
;    APPLICANT:  David Cooper
;    TITLE OF INVENTION:  NON-PATHOGENIC STRAINS OF HIV-1
;    NUMBER OF SEQUENCES:  841
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SCULLY, SCOTT, MURPHY & PRESSER
;      STREET:  400 GARDEN CITY PLAZA
;      CITY:  GARDEN CITY
;      STATE:  NEW YORK
;      COUNTRY:  U.S.A.
;      ZIP:  11530-0299
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/488,551B
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PM3864 (AU)
;      FILING DATE:  14-FEB-1994
```

```
;       APPLICATION NUMBER:  PM4002 (AU)
;       FILING DATE:  21-FEB-1994
;       APPLICATION NUMBER:  PN0284 (AU)
;       FILING DATE:  23-DEC-1994
;       APPLICATION NUMBER:  US 08/388,353
;       FILING DATE:  14-FEB-1995
;       APPLICATION NUMBER:  PN3021/95
;       FILING DATE:  17-MAY-1995
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  FRANK S. DIGIGLIO
;       REFERENCE/DOCKET NUMBER:  9606Z
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (516) 742-4343
;       TELEFAX:  (516) 742-4366
;   INFORMATION FOR SEQ ID NO:  504:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  10 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA
US-08-488-551B-504

  Query Match             62.0%;  Score 6.2;  DB 3;  Length 10;
  Best Local Similarity   40.0%;  Pred. No. 8.9e+04;
  Matches    4;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 UGCDGGHNMD 10
              :||:||: ::
Db         10 TGCAGGCTCA 1


RESULT 5
US-08-488-551B-822/c
; Sequence 822, Application US/08488551B
; Patent No. 6015661
;  GENERAL INFORMATION:
;    APPLICANT:  Nicholas J. Deacon
;    APPLICANT:  Dale A. McPhee
;    APPLICANT:  David Cooper
;    TITLE OF INVENTION:  NON-PATHOGENIC STRAINS OF HIV-1
;    NUMBER OF SEQUENCES:  841
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SCULLY, SCOTT, MURPHY & PRESSER
;      STREET:  400 GARDEN CITY PLAZA
;      CITY:  GARDEN CITY
;      STATE:  NEW YORK
;      COUNTRY:  U.S.A.
;      ZIP:  11530-0299
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/488,551B
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PM3864 (AU)
;      FILING DATE:  14-FEB-1994
;      APPLICATION NUMBER:  PM4002 (AU)
;      FILING DATE:  21-FEB-1994
;      APPLICATION NUMBER:  PN0284 (AU)
;      FILING DATE:  23-DEC-1994
;      APPLICATION NUMBER:  US 08/388,353
;      FILING DATE:  14-FEB-1995
;      APPLICATION NUMBER:  PN3021/95
;      FILING DATE:  17-MAY-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  FRANK S. DIGIGLIO
;      REFERENCE/DOCKET NUMBER:  9606Z
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (516) 742-4343
;      TELEFAX:  (516) 742-4366
;  INFORMATION FOR SEQ ID NO:  822:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
US-08-488-551B-822

  Query Match             62.0%;  Score 6.2;  DB 3;  Length 10;
  Best Local Similarity   40.0%;  Pred. No. 8.9e+04;
  Matches    4;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 UGCDGGHNMD 10
              :||:||: ::
Db         10 TGCAGGCTCA 1


RESULT 6
US-08-522-384-102
; Sequence 102, Application US/08522384
; Patent No. 6110667
; GENERAL INFORMATION:
;  APPLICANT: LOPEZ-NIETO, CARLOS E
;  APPLICANT: NIGAM, SANJAY KUMAR
;  TITLE OF INVENTION: PROCESSES, APPARATUS AND COMPOSITIONS FOR
;  TITLE OF INVENTION: CHARACTERIZING NUCLEOTIDE SEQUENCES
;  FILE REFERENCE: 2458-4029
;  CURRENT APPLICATION NUMBER: US/08/522,384
;  CURRENT FILING DATE: 1996-11-15
;  NUMBER OF SEQ ID NOS: 122
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 102
;   LENGTH: 10
;   TYPE: DNA
;   ORGANISM: Unknown Organism
;   FEATURE:
;   OTHER INFORMATION: Description of Unknown Organism:  Primer
US-08-522-384-102

  Query Match             62.0%;  Score 6.2;  DB 3;  Length 10;
  Best Local Similarity   40.0%;  Pred. No. 8.9e+04;
  Matches    4;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 UGCDGGHNMD 10
              :||:||: ::
Db          1 TGCTGGTCAT 10


RESULT 7
US-09-261-115-71
; Sequence 71, Application US/09261115
; Patent No. 6458584
; GENERAL INFORMATION:
;  APPLICANT: MIRZABEKOV, ANDREI
;  APPLICANT:  GUSCHIN, DMITRY Y.
;  APPLICANT:  SHIK, VALENTINE
;  APPLICANT:  DROBYSHEV, ALEKSEI
;  APPLICANT:  FOTIN, ALEXANDER
;  APPLICANT:  YERSHOV, GENNADIY
;  APPLICANT:  LYSOV, YU
;  TITLE OF INVENTION: CUSTOMIZED OLIGONUCLEOTIDE MICROCHIPS THAT CONVERT
;  TITLE OF INVENTION:  MULTIPLE GENETIC INFORMATION TO SIMPLE PATTERNS, ARE
;  TITLE OF INVENTION:  PORTABLE AND REUSABLE
;  FILE REFERENCE: 21416/90184
;  CURRENT APPLICATION NUMBER: US/09/261,115
;  CURRENT FILING DATE:  1999-03-03
;  NUMBER OF SEQ ID NOS: 78
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 71
;   LENGTH: 10
;   TYPE: DNA
```

```
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Customized
;   OTHER INFORMATION: oligonucleotide
US-09-261-115-71

  Query Match             62.0%;  Score 6.2;  DB 4;  Length 10;
  Best Local Similarity   40.0%;  Pred. No. 8.9e+04;
  Matches    4;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 UGCDGGHNMD 10
              :||:||: ::
Db          1 TGCGGGTCAA 10


RESULT 8
US-09-508-753B-22
; Sequence 22, Application US/09508753B
; Patent No. 6544736
; GENERAL INFORMATION:
;  APPLICANT: Akira SHIMAMOTO
;  APPLICANT:  Yasuhiro FURUICHI
;  APPLICANT:  Yuko SHIBATA
;  APPLICANT:  Hiroko FUNAKI
;  APPLICANT:  Eiji OHARA
;  APPLICANT:  Masanori WATAHIKI
;  TITLE OF INVENTION: Method for Synthesizing cDNA from mRNA sample
;  FILE REFERENCE: 00162/HG
;  CURRENT APPLICATION NUMBER: US/09/508,753B
;  CURRENT FILING DATE:  2000-06-16
;  PRIOR APPLICATION NUMBER: JP 9/270324
;  PRIOR FILING DATE: 1997-09-18
;  NUMBER OF SEQ ID NOS: 472
; SEQ ID NO 22
;   LENGTH: 10
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Primer
US-09-508-753B-22

  Query Match             62.0%;  Score 6.2;  DB 4;  Length 10;
  Best Local Similarity   40.0%;  Pred. No. 8.9e+04;
  Matches    4;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 UGCDGGHNMD 10
              :||:||: ::
Db          1 TGCTGGAAAG 10


RESULT 9
US-08-481-658B-75
; Sequence 75, Application US/08481658B
; Patent No. 5955075
;  GENERAL INFORMATION:
;    APPLICANT:  Zavada, Jan
;    APPLICANT:  Pastorekova, Silvia
;    APPLICANT:  Pastorek, Jaromir
;    TITLE OF INVENTION:  MN Gene and Protein
;    NUMBER OF SEQUENCES:  86
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Leona L. Lauder
;      STREET:  6 Mariposa Court
;      CITY:  Tiburon
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94920
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/481,658B
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/260,190
;      FILING DATE:  15-JUN-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lauder, Leona L.
;      REGISTRATION NUMBER:  30,863
;      REFERENCE/DOCKET NUMBER:  D-0021.3E
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-435-2034
;      TELEFAX:  415-435-0727
;  INFORMATION FOR SEQ ID NO:  75:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  11 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;      DESCRIPTION:  5' donor consensus splice sequence
US-08-481-658B-75

  Query Match             62.0%;  Score 6.2;  DB 2;  Length 11;
  Best Local Similarity   40.0%;  Pred. No. 8.9e+04;
  Matches    4;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 UGCDGGHNMD 10
              :||:||: ::
Db          1 TGCTGGTGAG 10


RESULT 10
US-08-477-504A-75
; Sequence 75, Application US/08477504A
; Patent No. 5972353
;  GENERAL INFORMATION:
;    APPLICANT:  Zavada, Jan
;    APPLICANT:  Pastorekova, Silvia
;    APPLICANT:  Pastorek, Jaromir
;    TITLE OF INVENTION:  MN Gene and Protein
;    NUMBER OF SEQUENCES:  86
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Leona L. Lauder
;      STREET:  6 Mariposa Court
;      CITY:  Tiburon
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94920
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/477,504A
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/260,190
;      FILING DATE:  15-JUN-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lauder, Leona L.
;      REGISTRATION NUMBER:  30,863
;      REFERENCE/DOCKET NUMBER:  D-0021.3D
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-435-2034
;      TELEFAX:  415-435-0727
;  INFORMATION FOR SEQ ID NO:  75:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  11 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
```

```
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;      DESCRIPTION:  5' donor consensus splice sequence
US-08-477-504A-75

  Query Match             62.0%;  Score 6.2;  DB 2;  Length 11;
  Best Local Similarity   40.0%;  Pred. No. 8.9e+04;
  Matches    4;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 UGCDGGHNMD 10
              :||:||: ::
Db          1 TGCTGGTGAG 10


RESULT 11
US-08-486-756A-75
; Sequence 75, Application US/08486756A
; Patent No. 5981711
;  GENERAL INFORMATION:
;    APPLICANT:  Zavada, Jan
;    APPLICANT:  Pastorekova, Silvia
;    APPLICANT:  Pastorek, Jaromir
;    TITLE OF INVENTION:  MN Gene and Protein
;    NUMBER OF SEQUENCES:  86
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Leona L. Lauder
;      STREET:  6 Mariposa Court
;      CITY:  Tiburon
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94920
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/486,756A
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/260,190
;      FILING DATE:  15-JUN-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lauder, Leona L.
;      REGISTRATION NUMBER:  30,863
;      REFERENCE/DOCKET NUMBER:  D-0021.3C
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-435-2034
;      TELEFAX:  415-435-0727
;  INFORMATION FOR SEQ ID NO:  75:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  11 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;      DESCRIPTION:  5' donor consensus splice sequence
US-08-486-756A-75

  Query Match             62.0%;  Score 6.2;  DB 2;  Length 11;
  Best Local Similarity   40.0%;  Pred. No. 8.9e+04;
  Matches    4;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 UGCDGGHNMD 10
              :||:||: ::
Db          1 TGCTGGTGAG 10


RESULT 12
US-08-485-862B-75
; Sequence 75, Application US/08485862B
; Patent No. 5989838
;  GENERAL INFORMATION:
;    APPLICANT:  Zavada, Jan
;    APPLICANT:  Pastorekova, Silvia
;    APPLICANT:  Pastorek, Jaromir
;    TITLE OF INVENTION:  MN Gene and Protein
;    NUMBER OF SEQUENCES:  86
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Leona L. Lauder
;      STREET:  6 Mariposa Court
;      CITY:  Tiburon
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94920
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/485,862B
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/477,504
;      FILING DATE:  07-JUN-1995
;      APPLICATION NUMBER:  US 08/260,190
;      FILING DATE:  15-JUN-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lauder, Leona L.
;      REGISTRATION NUMBER:  30,863
;      REFERENCE/DOCKET NUMBER:  D-0021.3D
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-435-2034
;      TELEFAX:  415-435-0727
;  INFORMATION FOR SEQ ID NO:  75:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  11 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;      DESCRIPTION:  5' donor consensus splice sequence
US-08-485-862B-75

  Query Match             62.0%;  Score 6.2;  DB 2;  Length 11;
  Best Local Similarity   40.0%;  Pred. No. 8.9e+04;
  Matches    4;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 UGCDGGHNMD 10
              :||:||: ::
Db          1 TGCTGGTGAG 10


RESULT 13
US-08-787-739-75
; Sequence 75, Application US/08787739
; Patent No. 6027887
;  GENERAL INFORMATION:
;    APPLICANT:  Zavada, Jan
;    APPLICANT:  Pastorekova, Silvia
;    APPLICANT:  Pastorek, Jaromir
;    TITLE OF INVENTION:  MN Gene and Protein
;    NUMBER OF SEQUENCES:  96
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Leona L. Lauder
;      STREET:  369 Pine Street, Suite 610
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
```

```
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30 (EPO)
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/787,739
;       FILING DATE:  24-JAN-1997
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/485,049
;       FILING DATE:  07-JUN-1995
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/486,756
;       FILING DATE:  07-JUN-1995
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/477,504
;       FILING DATE:  07-JUN-1995
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/481,658
;       FILING DATE:  07-JUN-1995
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/485,862
;       FILING DATE:  07-JUN-1995
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/485,863
;       FILING DATE:  07-JUN-1995
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/487,077
;       FILING DATE:  07-JUN-1995
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Lauder, Leona L.
;       REGISTRATION NUMBER:  30,863
;       REFERENCE/DOCKET NUMBER:  D-0021.4
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  415-981-2034
;       TELEFAX:  415-981-0332
;   INFORMATION FOR SEQ ID NO:  75:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  11 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  double
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
;       DESCRIPTION:  5' donor consensus splice sequence
US-08-787-739-75

  Query Match             62.0%;  Score 6.2;  DB 3;  Length 11;
  Best Local Similarity   40.0%;  Pred. No. 8.9e+04;
  Matches    4;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 UGCDGGHNMD 10
              :||:||: ::
Db          1 TGCTGGTGAG 10


RESULT 14
US-08-487-077A-75
; Sequence 75, Application US/08487077A
; Patent No. 6069242
;  GENERAL INFORMATION:
;    APPLICANT:  Zavada, Jan
;    APPLICANT:  Pastorekova, Silvia
;    APPLICANT:  Pastorek, Jaromir
;    TITLE OF INVENTION:  MN Gene and Protein
;    NUMBER OF SEQUENCES:  86
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Leona L. Lauder
;      STREET:  6 Mariposa Court
;      CITY:  Tiburon
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94920
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30 (EPO)
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/487,077A
;       FILING DATE:  07-JUN-1995
;       CLASSIFICATION:  514
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/260,190
;       FILING DATE:  15-JUN-1994
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Lauder, Leona L.
;       REGISTRATION NUMBER:  30,863
;       REFERENCE/DOCKET NUMBER:  D-0021.3H
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  415-435-2034
;       TELEFAX:  415-435-0727
;   INFORMATION FOR SEQ ID NO:  75:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  11 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
;       DESCRIPTION:  5' donor consensus splice sequence
US-08-487-077A-75

  Query Match             62.0%;  Score 6.2;  DB 3;  Length 11;
  Best Local Similarity   40.0%;  Pred. No. 8.9e+04;
  Matches    4;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 UGCDGGHNMD 10
              :||:||: ::
Db          1 TGCTGGTGAG 10


RESULT 15
US-08-485-863A-75
; Sequence 75, Application US/08485863A
; Patent No. 6093548
;  GENERAL INFORMATION:
;    APPLICANT:  Zavada, Jan
;    APPLICANT:  Pastorekova, Silvia
;    APPLICANT:  Pastorek, Jaromir
;    TITLE OF INVENTION:  MN Gene and Protein
;    NUMBER OF SEQUENCES:  86
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Leona L. Lauder
;      STREET:  6 Mariposa Court
;      CITY:  Tiburon
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94920
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/485,863A
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/260,190
;      FILING DATE:  15-JUN-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lauder, Leona L.
;      REGISTRATION NUMBER:  30,863
;      REFERENCE/DOCKET NUMBER:  D-0021.3G
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-435-2034
;      TELEFAX:  415-435-0727
;  INFORMATION FOR SEQ ID NO:  75:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  11 base pairs
```

```
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;      DESCRIPTION:  5' donor consensus splice sequence
US-08-485-863A-75

  Query Match             62.0%;  Score 6.2;  DB 3;  Length 11;
  Best Local Similarity   40.0%;  Pred. No. 8.9e+04;
  Matches    4;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 UGCDGGHNMD 10
              :||:||: ::
Db          1 TGCTGGTGAG 10


Search completed: October 19, 2004, 16:08:36
Job time : 58 secs
```

This Page Blank (uspto)

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: October 19, 2004, 15:00:01 ; Search time 260 Seconds
(without alignments)
196.753 Million cell updates/sec

Title: US-09-762-648-4
Perfect score: 10
Sequence: 1 UGCDGGHNMD 10

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3403857 seqs, 2557783690 residues

Total number of hits satisfying chosen parameters: 6807714

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
20: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
c 1	6.2	62.0	10	13	US-10-033-145-59	Sequence 59, Appl
c 2	6.2	62.0	10	13	US-10-033-145-312	Sequence 312, App
3	6.2	62.0	10	13	US-10-084-700-17	Sequence 17, Appl
4	6.2	62.0	10	15	US-10-212-476-71	Sequence 71, Appl
c 5	6.2	62.0	11	9	US-09-005-243-99	Sequence 99, Appl
c 6	6.2	62.0	11	9	US-09-224-683-99	Sequence 99, Appl
7	6.2	62.0	11	9	US-09-772-719-75	Sequence 75, Appl
c 8	6.2	62.0	11	10	US-09-918-715-162	Sequence 162, App
9	6.2	62.0	11	10	US-09-967-237-75	Sequence 75, Appl
c 10	6.2	62.0	11	15	US-10-314-322-303	Sequence 303, App
c 11	6.2	62.0	11	17	US-10-203-969A-5	Sequence 5, Appli
12	6.2	62.0	11	17	US-10-450-797-357	Sequence 357, App
c 13	6.2	62.0	11	17	US-10-450-797-421	Sequence 421, App
c 14	6.2	62.0	11	17	US-10-450-797-663	Sequence 663, App
c 15	6.2	62.0	11	17	US-10-175-608-99	Sequence 99, Appl
c 16	6.2	62.0	12	13	US-10-085-108-6	Sequence 6, Appli
17	6.2	62.0	12	15	US-10-076-047A-110	Sequence 110, App
c 18	6.2	62.0	12	15	US-10-160-237-6	Sequence 6, Appli
c 19	6.2	62.0	13	10	US-09-823-181-4	Sequence 4, Appli
c 20	6.2	62.0	14	8	US-08-980-068B-12	Sequence 12, Appl
c 21	6.2	62.0	14	8	US-08-980-068B-15	Sequence 15, Appl
c 22	6.2	62.0	14	8	US-08-980-068B-17	Sequence 17, Appl
c 23	6.2	62.0	14	10	US-09-880-313A-131	Sequence 131, App
c 24	6.2	62.0	14	15	US-10-084-839-3065	Sequence 3065, Ap
c 25	6.2	62.0	14	15	US-10-084-839-3069	Sequence 3069, Ap
26	6.2	62.0	14	15	US-10-091-281-346	Sequence 346, App
27	6.2	62.0	14	17	US-10-318-390-11	Sequence 11, Appl
c 28	6.2	62.0	15	9	US-09-504-231A-1160	Sequence 1160, Ap
c 29	6.2	62.0	15	9	US-09-504-231A-1161	Sequence 1161, Ap
30	6.2	62.0	15	9	US-09-152-059-11	Sequence 11, Appl
31	6.2	62.0	15	9	US-09-152-059-13	Sequence 13, Appl
32	6.2	62.0	15	9	US-09-152-059-20	Sequence 20, Appl
33	6.2	62.0	15	9	US-09-152-059-21	Sequence 21, Appl
c 34	6.2	62.0	15	9	US-09-274-553D-1160	Sequence 1160, Ap
c 35	6.2	62.0	15	9	US-09-274-553D-1161	Sequence 1161, Ap
36	6.2	62.0	15	9	US-09-544-934B-106	Sequence 106, App
c 37	6.2	62.0	15	9	US-09-263-959-59	Sequence 59, Appl
38	6.2	62.0	15	9	US-09-263-959-69	Sequence 69, Appl
c 39	6.2	62.0	15	9	US-09-263-959-83	Sequence 83, Appl
c 40	6.2	62.0	15	10	US-09-771-933-175	Sequence 175, App
c 41	6.2	62.0	15	10	US-09-880-313A-19	Sequence 19, Appl
c 42	6.2	62.0	15	10	US-09-880-313A-27	Sequence 27, Appl
c 43	6.2	62.0	15	10	US-09-880-313A-29	Sequence 29, Appl
c 44	6.2	62.0	15	10	US-09-880-313A-49	Sequence 49, Appl
c 45	6.2	62.0	15	10	US-09-880-313A-51	Sequence 51, Appl

ALIGNMENTS

RESULT 1
US-10-033-145-59/c
; Sequence 59, Application US/10033145
; Publication No. US20020151515A1
; GENERAL INFORMATION:
; APPLICANT: GENZYME CORPORATION
; APPLICANT: ROBERTS, BRUCE
; APPLICANT: SHANKARA, SRINIVAS
; TITLE OF INVENTION: PREPARATION AND USE OF SUPERIOR VACCINES
; FILE REFERENCE: GA0201C
; CURRENT APPLICATION NUMBER: US/10/033,145
; CURRENT FILING DATE: 2001-11-05
; PRIOR APPLICATION NUMBER: PCT/US99/13800
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 2137
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 59
; LENGTH: 10
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-033-145-59

Query Match 62.0%; Score 6.2; DB 13; Length 10;
Best Local Similarity 40.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 UGCDGGHNMD 10
              :||:||: ::
Db         10 TGCTGGCTAG 1
```

RESULT 2
US-10-033-145-312/c
; Sequence 312, Application US/10033145
; Publication No. US20020151515A1
; GENERAL INFORMATION:

```
;  APPLICANT: GENZYME CORPORATION
;  APPLICANT:  ROBERTS, BRUCE
;  APPLICANT:  SHANKARA, SRINIVAS
;  TITLE OF INVENTION: PREPARATION AND USE OF SUPERIOR VACCINES
;  FILE REFERENCE: GA0201C
;  CURRENT APPLICATION NUMBER: US/10/033,145
;  CURRENT FILING DATE:  2001-11-05
;  PRIOR APPLICATION NUMBER: PCT/US99/13800
;  PRIOR FILING DATE: 1999-06-18
;  NUMBER OF SEQ ID NOS: 2137
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 312
;   LENGTH: 10
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-033-145-312

  Query Match             62.0%;  Score 6.2;  DB 13;  Length 10;
  Best Local Similarity   40.0%;  Pred. No. 7.8e+05;
  Matches    4;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 UGCDGGHNMD 10
              :||:||: ::
Db         10 TGCAGGAAAA 1


RESULT 3
US-10-084-700-17
; Sequence 17, Application US/10084700
; Publication No. US20020160403A1
; GENERAL INFORMATION:
;  APPLICANT: Seeley, Todd
;  TITLE OF INVENTION: huBUB3 GENE INVOLVED IN HUMAN CANCERS
;  FILE REFERENCE: PP-01406.004/200130.438D1
;  CURRENT APPLICATION NUMBER: US/10/084,700
;  CURRENT FILING DATE:  2002-02-27
;  NUMBER OF SEQ ID NOS: 32
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17
;   LENGTH: 10
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-084-700-17

  Query Match             62.0%;  Score 6.2;  DB 13;  Length 10;
  Best Local Similarity   40.0%;  Pred. No. 7.8e+05;
  Matches    4;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 UGCDGGHNMD 10
              :||:||: ::
Db          1 TGCAGGTCCA 10


RESULT 4
US-10-212-476-71
; Sequence 71, Application US/10212476
; Publication No. US20030157509A1
; GENERAL INFORMATION:
;  APPLICANT: MIRZABEKOV, ANDREI
;  APPLICANT:  GUSCHIN, DMITRY Y.
;  APPLICANT:  SHIK, VALENTINE
;  APPLICANT:  DROBYSHEV, ALEKSEI
;  APPLICANT:  FOTIN, ALEXANDER
;  APPLICANT:  YERSHOV, GENNADIY
;  APPLICANT:  LYSOV, YU
;  TITLE OF INVENTION: CUSTOMIZED OLIGONUCLEOTIDE MICROCHIPS THAT CONVERT
;  TITLE OF INVENTION:  MULTIPLE GENETIC INFORMATION TO SIMPLE PATTERNS, ARE
;  TITLE OF INVENTION:  PORTABLE AND REUSABLE
;  FILE REFERENCE: 21416/90184
;  CURRENT APPLICATION NUMBER: US/10/212,476
;  CURRENT FILING DATE:  2002-08-08
;  PRIOR APPLICATION NUMBER: US/09/261,115
;  PRIOR FILING DATE: 1999-03-03
;  NUMBER OF SEQ ID NOS: 78
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 71
;   LENGTH: 10
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Customized
;   OTHER INFORMATION: oligonucleotide
US-10-212-476-71

  Query Match             62.0%;  Score 6.2;  DB 15;  Length 10;
  Best Local Similarity   40.0%;  Pred. No. 7.8e+05;
  Matches    4;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 UGCDGGHNMD 10
              :||:||: ::
Db          1 TGCGGGTCAA 10


RESULT 5
US-09-005-243-99/c
; Sequence 99, Application US/09005243
; Patent No. US20020018763A1
;  GENERAL INFORMATION:
;    APPLICANT:  Zsebo, Krisztina M.
;    APPLICANT:  Bosselman, Robert A.
;    APPLICANT:  Suggs, Sidney V.
;    APPLICANT:  Martin, Francis H.
;    TITLE OF INVENTION:  Stem Cell Factor
;    NUMBER OF SEQUENCES:  104
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/005,243
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/449,653
;      FILING DATE:  24-MAY-1995
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/982,255
;      FILING DATE:  25-NOV-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/589,701
;      FILING DATE:  01-OCT-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/573,616
;      FILING DATE:  24-AUG-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/537,198
;      FILING DATE:  11-JUN-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/422,383
;      FILING DATE:  16-OCT-1989
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Clough, David W.
;      REGISTRATION NUMBER:  36,107
;      REFERENCE/DOCKET NUMBER:  01017/34465
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
```

```
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  99:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  11 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
US-09-005-243-99

  Query Match             62.0%;  Score 6.2;  DB 9;  Length 11;
  Best Local Similarity   40.0%;  Pred. No. 7.7e+05;
  Matches    4;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 UGCDGGHNMD 10
              :||:||: ::
Db         11 TGCAGGAGAT 2


RESULT 6
US-09-224-683-99/c
; Sequence 99, Application US/09224683
; Patent No. US20020031491A1
;  GENERAL INFORMATION:
;    APPLICANT:  Zsebo, Krisztina M.
;    APPLICANT:  Bosselman, Robert A.
;    APPLICANT:  Suggs, Sidney V.
;    APPLICANT:  Martin, Francis H.
;    TITLE OF INVENTION:  Stem Cell Factor: Composition Claims
;    NUMBER OF SEQUENCES:  104
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/224,683
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  09/005,893
;      FILING DATE:  12-JAN-1998
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/449,653
;      FILING DATE:  24-MAY-1995
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/982,255
;      FILING DATE:  25-NOV-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/589,701
;      FILING DATE:  01-OCT-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/573,616
;      FILING DATE:  24-AUG-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/537,198
;      FILING DATE:  11-JUN-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/422,383
;      FILING DATE:  16-OCT-1989
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Clough, David W.
;      REGISTRATION NUMBER:  36,107
;      REFERENCE/DOCKET NUMBER:  01017/35136
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  99:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  11 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
US-09-224-683-99

  Query Match             62.0%;  Score 6.2;  DB 9;  Length 11;
  Best Local Similarity   40.0%;  Pred. No. 7.7e+05;
  Matches    4;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 UGCDGGHNMD 10
              :||:||: ::
Db         11 TGCAGGAGAT 2


RESULT 7
US-09-772-719-75
; Sequence 75, Application US/09772719
; Patent No. US20020137910A1
;  GENERAL INFORMATION:
;    APPLICANT:  Zavada, Jan
;    APPLICANT:  Pastorekova, Silvia
;    APPLICANT:  Pastorek, Jaromir
;    TITLE OF INVENTION:  MN Gene and Protein
;    NUMBER OF SEQUENCES:  86
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Leona L. Lauder
;      STREET:  369 Pine Street
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/772,719
;      FILING DATE:  30-JAN-2001
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/485,049
;      FILING DATE:  07-JUN-1995
;      APPLICATION NUMBER:  US 08/260,190
;      FILING DATE:  15-JUN-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lauder, Leona L.
;      REGISTRATION NUMBER:  30,863
;      REFERENCE/DOCKET NUMBER:  D-0021.3E
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-981-2034
;      TELEFAX:  415-981-0332
;  INFORMATION FOR SEQ ID NO:  75:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  11 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;      DESCRIPTION:  5' donor consensus splice sequence
US-09-772-719-75

  Query Match             62.0%;  Score 6.2;  DB 9;  Length 11;
  Best Local Similarity   40.0%;  Pred. No. 7.7e+05;
  Matches    4;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;
```

```
Qy          1 UGCDGGHNMD 10
              :||:||: ::
Db          1 TGCTGGTGAG 10


RESULT 8
US-09-918-715-162/c
; Sequence 162, Application US/09918715
; Publication No. US20030017157A1
; GENERAL INFORMATION:
;  APPLICANT: Brad St. Croix
;  APPLICANT:  Bert Vogelstein
;  APPLICANT:  Kenneth Kinzler
;  TITLE OF INVENTION: ENDOTHELIAL CELL EXPRESSION PATTERNS
;  FILE REFERENCE: 1107.00134
;  CURRENT APPLICATION NUMBER: US/09/918,715
;  CURRENT FILING DATE:  2001-08-01
;  PRIOR APPLICATION NUMBER: 60/222,599
;  PRIOR FILING DATE: 2000-08-02
;  PRIOR APPLICATION NUMBER: 60/224,360
;  PRIOR FILING DATE: 2000-08-11
;  PRIOR APPLICATION NUMBER: 60/282,850
;  PRIOR FILING DATE: 2000-04-11
;  NUMBER OF SEQ ID NOS: 358
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 162
;   LENGTH: 11
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-918-715-162

  Query Match             62.0%;  Score 6.2;  DB 10;  Length 11;
  Best Local Similarity   40.0%;  Pred. No. 7.7e+05;
  Matches    4;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 UGCDGGHNMD 10
              :||:||: ::
Db         10 TGCAGGTTAT 1


RESULT 9
US-09-967-237-75
; Sequence 75, Application US/09967237
; Publication No. US20030049828A1
; GENERAL INFORMATION:
;  APPLICANT: Zavada, Jan
;  APPLICANT:  Pastorekova, Silvia
;  APPLICANT:  Pastorek, Jaromir
;  TITLE OF INVENTION: MN Gene and Protein
;  FILE REFERENCE: D-0021.5B-2
;  CURRENT APPLICATION NUMBER: US/09/967,237
;  CURRENT FILING DATE:  2001-09-27
;  PRIOR APPLICATION NUMBER: 09/178,115
;  PRIOR FILING DATE: 1998-10-23
;  NUMBER OF SEQ ID NOS: 116
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 75
;   LENGTH: 11
;   TYPE: DNA
;   ORGANISM: HUMAN
US-09-967-237-75

  Query Match             62.0%;  Score 6.2;  DB 10;  Length 11;
  Best Local Similarity   40.0%;  Pred. No. 7.7e+05;
  Matches    4;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 UGCDGGHNMD 10
              :||:||: ::
Db          1 TGCTGGTGAG 10


RESULT 10
US-10-314-322-303/c
; Sequence 303, Application US/10314322
; Publication No. US20030229911A1
; GENERAL INFORMATION:
;  APPLICANT: Heber-Katz, Ellen
;  TITLE OF INVENTION: Compositions and Methods for Wound
;  TITLE OF INVENTION:  Healing
;  FILE REFERENCE: 000486.00016
;  CURRENT APPLICATION NUMBER: US/10/314,322
;  CURRENT FILING DATE:  2002-12-09
;  PRIOR APPLICATION NUMBER: US 60/074,737
;  PRIOR FILING DATE: 1998-02-13
;  PRIOR APPLICATION NUMBER: US 60/097,937
;  PRIOR FILING DATE: 1998-08-26
;  PRIOR APPLICATION NUMBER: US 60/102,051
;  PRIOR FILING DATE: 1998-09-28
;  PRIOR APPLICATION NUMBER: US 09/249,155
;  PRIOR FILING DATE: 1999-02-12
;  NUMBER OF SEQ ID NOS: 346
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 303
;   LENGTH: 11
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-10-314-322-303

  Query Match             62.0%;  Score 6.2;  DB 15;  Length 11;
  Best Local Similarity   40.0%;  Pred. No. 7.7e+05;
  Matches    4;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 UGCDGGHNMD 10
              :||:||: ::
Db         11 TGCTGGAGAG 2


RESULT 11
US-10-203-969A-5/c
; Sequence 5, Application US/10203969A
; Publication No. US20040110224A1
; GENERAL INFORMATION:
;  APPLICANT: Puijk, Wouter C.
;  APPLICANT:  Dijk van, Evert
;  APPLICANT:  Slootstra, Jelle W.
;  TITLE OF INVENTION: Segment synthesis
;  FILE REFERENCE: P50200US00
;  CURRENT APPLICATION NUMBER: US/10/203,969A
;  CURRENT FILING DATE:  2003-07-07
;  PRIOR APPLICATION NUMBER: EP 00200536.1
;  PRIOR FILING DATE: 2000-02-16
;  PRIOR APPLICATION NUMBER: PCT/NL01/00131
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 660
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
;   LENGTH: 11
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: unspecified
;   OTHER INFORMATION: biotinylated DNA
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(11)
US-10-203-969A-5

  Query Match             62.0%;  Score 6.2;  DB 17;  Length 11;
  Best Local Similarity   40.0%;  Pred. No. 7.7e+05;
  Matches    4;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 UGCDGGHNMD 10
              :||:||: ::
Db         10 TGCGGGTGCG 1
```

```
RESULT 12
US-10-450-797-357
; Sequence 357, Application US/10450797
; Publication No. US20040142335A1
; GENERAL INFORMATION:
;  APPLICANT: Petersohn, Dirk
;  APPLICANT:  Conradt, Marcus
;  APPLICANT:  Hofmann, Kay
;  TITLE OF INVENTION: METHOD FOR DETERMINING SKIN STRESS OR SKIN AGEING IN VITRO
;  FILE REFERENCE: HENK-0041
;  CURRENT APPLICATION NUMBER: US/10/450,797
;  CURRENT FILING DATE:  2003-12-04
;  PRIOR APPLICATION NUMBER: PCT/EP01/15178
;  PRIOR FILING DATE: 2001-12-20
;  PRIOR APPLICATION NUMBER: DE 101 00 121.5
;  PRIOR FILING DATE: 2001-01-03
;  NUMBER OF SEQ ID NOS: 1435
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 357
;   LENGTH: 11
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-450-797-357

  Query Match             62.0%;  Score 6.2;  DB 17;  Length 11;
  Best Local Similarity   40.0%;  Pred. No. 7.7e+05;
  Matches    4;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 UGCDGGHNMD 10
              :||:||: ::
Db          2 TGCTGGAGAA 11


RESULT 13
US-10-450-797-421/c
; Sequence 421, Application US/10450797
; Publication No. US20040142335A1
; GENERAL INFORMATION:
;  APPLICANT: Petersohn, Dirk
;  APPLICANT:  Conradt, Marcus
;  APPLICANT:  Hofmann, Kay
;  TITLE OF INVENTION: METHOD FOR DETERMINING SKIN STRESS OR SKIN AGEING IN VITRO
;  FILE REFERENCE: HENK-0041
;  CURRENT APPLICATION NUMBER: US/10/450,797
;  CURRENT FILING DATE:  2003-12-04
;  PRIOR APPLICATION NUMBER: PCT/EP01/15178
;  PRIOR FILING DATE: 2001-12-20
;  PRIOR APPLICATION NUMBER: DE 101 00 121.5
;  PRIOR FILING DATE: 2001-01-03
;  NUMBER OF SEQ ID NOS: 1435
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 421
;   LENGTH: 11
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-450-797-421

  Query Match             62.0%;  Score 6.2;  DB 17;  Length 11;
  Best Local Similarity   40.0%;  Pred. No. 7.7e+05;
  Matches    4;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 UGCDGGHNMD 10
              :||:||: ::
Db         10 TGCGGGTCAT 1


RESULT 14
US-10-450-797-663/c
; Sequence 663, Application US/10450797
; Publication No. US20040142335A1
; GENERAL INFORMATION:
;  APPLICANT: Petersohn, Dirk
;  APPLICANT:  Conradt, Marcus
;  APPLICANT:  Hofmann, Kay
;  TITLE OF INVENTION: METHOD FOR DETERMINING SKIN STRESS OR SKIN AGEING IN VITRO
;  FILE REFERENCE: HENK-0041
;  CURRENT APPLICATION NUMBER: US/10/450,797
;  CURRENT FILING DATE:  2003-12-04
;  PRIOR APPLICATION NUMBER: PCT/EP01/15178
;  PRIOR FILING DATE: 2001-12-20
;  PRIOR APPLICATION NUMBER: DE 101 00 121.5
;  PRIOR FILING DATE: 2001-01-03
;  NUMBER OF SEQ ID NOS: 1435
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 663
;   LENGTH: 11
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-450-797-663

  Query Match             62.0%;  Score 6.2;  DB 17;  Length 11;
  Best Local Similarity   40.0%;  Pred. No. 7.7e+05;
  Matches    4;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 UGCDGGHNMD 10
              :||:||: ::
Db         11 TGCAGGACAG 2


RESULT 15
US-10-175-608-99/c
; Sequence 99, Application US/10175608
; Publication No. US20040181044A1
;  GENERAL INFORMATION:
;       APPLICANT: Zsebo, Krisztina M.
;                  Bosselman, Robert A.
;                  Suggs, Sidney V.
;                  Martin, Francis H.
;       TITLE OF INVENTION: Stem Cell Factor
;       NUMBER OF SEQUENCES: 104
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
;            STREET: 6300 Sears Tower, 233 South Wacker Drive
;            CITY: Chicago
;            STATE: Illinois
;            COUNTRY: United States of America
;            ZIP: 60606-6402
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/10/175,608
;            FILING DATE: 16-Oct-2002
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 09/635,249
;            FILING DATE: 07-AUG-2000
;            APPLICATION NUMBER: 09/486,546
;            FILING DATE: 24-MAY-1995
;            APPLICATION NUMBER: 08/172,329
;            FILING DATE: 21-DEC-1993
;            APPLICATION NUMBER: 07/982,255
;            FILING DATE: 25-NOV-1992
;            APPLICATION NUMBER: 07/684,535
;            FILING DATE: 10-APR-1991
;            APPLICATION NUMBER: 09/589,701
;            FILING DATE: 10-OCT-1991
;            APPLICATION NUMBER: 07/573,616
;            FILING DATE: 24-AUG-1990
;            APPLICATION NUMBER: 07/537,198
;            FILING DATE: 11-JUN-1990
;            APPLICATION NUMBER: 07/422,383
;            FILING DATE: 16-OCT-1989
```

```
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Clough, David W.
;            REGISTRATION NUMBER: 36,107
;            REFERENCE/DOCKET NUMBER: 01017/35199
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 312/474-6300
;            TELEFAX: 312/474-0448
;            TELEX: <Unknown>
;  INFORMATION FOR SEQ ID NO: 99:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 11 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       MOLECULE TYPE: DNA
;       SEQUENCE DESCRIPTION: SEQ ID NO: 99:
US-10-175-608-99

  Query Match             62.0%;  Score 6.2;  DB 17;  Length 11;
  Best Local Similarity   40.0%;  Pred. No. 7.7e+05;
  Matches    4;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 UGCDGGHNMD 10
              :||:||: ::
Db         11 TGCAGGAGAT 2


Search completed: October 19, 2004, 16:13:07
Job time : 261 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: October 19, 2004, 14:44:01 ; Search time 2180 Seconds
(without alignments)
167.155 Million cell updates/sec

Title: US-09-762-648-4
Perfect score: 10
Sequence: 1 UGCDGGHNMD 10

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_htc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	6.2	62.0	10	9	CL437053	PST4390-N
	2	6.2	62.0	13	9	AJ594173	Arabidops
c	3	6.2	62.0	19	8	AZ422762	AZ422762 1M0201P12
	4	6.2	62.0	19	8	AZ791433	AZ791433 2M0041A24
	5	6.2	62.0	19	8	AZ800056	AZ800056 2M0057E20
	6	6.2	62.0	19	9	CL435787	CL435787 PST1409-N
	7	6.2	62.0	20	8	AZ308311	AZ308311 1M0011J12
c	8	6.2	62.0	20	8	AZ590589	AZ590589 1M0400G15
	9	6.2	62.0	20	8	AZ630221	AZ630221 1M0483K12
	10	6.2	62.0	20	8	AZ789903	AZ789903 2M0038F15
c	11	6.2	62.0	21	7	CO790328	CO790328 NT009B_E1
	12	6.2	62.0	21	8	AZ594960	AZ594960 1M0407E16
	13	6.2	62.0	21	8	AZ776814	AZ776814 2M0010I17
	14	6.2	62.0	21	8	AZ819369	AZ819369 2M0089O21
c	15	6.2	62.0	22	8	AZ307716	AZ307716 1M0009N02
c	16	6.2	62.0	22	8	AZ309912	AZ309912 1M0017P15
	17	6.2	62.0	22	8	AZ416988	AZ416988 1M0192P23
	18	6.2	62.0	22	8	AZ501345	AZ501345 1M0340I11
	19	6.2	62.0	22	8	AZ610074	AZ610074 1M0435J11
	20	6.2	62.0	22	8	AZ822589	AZ822589 2M0095P21
	21	6.2	62.0	23	4	BG927951	BG927951 HNC45-1-F
c	22	6.2	62.0	23	8	AZ472899	AZ472899 1M0288L22
	23	6.2	62.0	23	8	AZ632990	AZ632990 1M0487M19
	24	6.2	62.0	23	8	AZ810074	AZ810074 2M0074J19
c	25	6.2	62.0	23	8	AZ977634	AZ977634 2M0253C10
	26	6.2	62.0	23	9	TA121C07P	AL465730 T. brucei
	27	6.2	62.0	23	9	TA278H12P	AL485533 T. brucei
c	28	6.2	62.0	23	9	CL670300	CL670300 PRI0161d_
	29	6.2	62.0	23	9	AG195455	AG195455 Pan trogl
	30	6.2	62.0	24	1	AL045569	AL045569 DKFZp434M
c	31	6.2	62.0	24	5	BQ589506	BQ589506 E012561-0
c	32	6.2	62.0	24	8	AZ308017	AZ308017 1M0010M05
c	33	6.2	62.0	24	8	AZ492799	AZ492799 1M0327B10
c	34	6.2	62.0	24	8	AZ505865	AZ505865 1M0346C18
	35	6.2	62.0	24	8	AZ821309	AZ821309 2M0094G05
	36	6.2	62.0	24	9	TA120A03P	AL462474 T. brucei
	37	6.2	62.0	24	9	AG187929	AG187929 Pan trogl
	38	6.2	62.0	25	1	AA880566	AA880566 vx41c05.r
	39	6.2	62.0	25	1	AA923337	AA923337 ol44e07.s
c	40	6.2	62.0	25	1	AA974358	AA974358 oq14f08.s
	41	6.2	62.0	25	5	BQ594927	BQ594927 S015257-0
c	42	6.2	62.0	25	8	AZ414090	AZ414090 1M0188F21
c	43	6.2	62.0	25	8	AZ510410	AZ510410 1M0354H21
	44	6.2	62.0	25	8	AZ769533	AZ769533 1M0570D22
c	45	6.2	62.0	25	9	AJ592851	AJ592851 Arabidops

ALIGNMENTS

RESULT 1
CL437053/c
LOCUS CL437053 10 bp DNA linear GSS 18-MAR-2004
DEFINITION PST4390-NR.Seq MICB1 Mus musculus genomic clone PST4390-NR.Seq, genomic survey sequence.
ACCESSION CL437053
VERSION CL437053.1 GI:45572462
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 10)
AUTHORS Hicks,G.G.
TITLE www.EScells.ca
JOURNAL Unpublished (2002)
COMMENT Contact: Hicks GG
Mammalian Functional Genomics Centre
Manitoba Institute of Cell Biology, Universitry of Manitoba
ON5029, 675 McDermot Ave, Winnipeg, MB R3E 0V9, Canada
Tel: 204 787 2133
Fax: 204 787 2190
Email: hicksgg@cc.umanitoba.ca
U3NeoSV1 gene trap. Tag generated by plasmid rescue. Additional sequence information and target gene cloning can be generated. ES cell line harboring insertion mutation of target gene is available. Sequence analysis available from
http://140.193.242.7/esdb/public_search_frame.php?PST=PST4390-NR.Seq
Class: Gene Trap.
FEATURES Location/Qualifiers
source 1. .10
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="129 sv"
/db_xref="taxon:10090"
/clone="PST4390-NR.Seq"
/sex="Male"
/cell_type="Embryonic stem cell"
/cell_line="D3H (J1 subclone)"
/clone_lib="MICB1"
/note="Vector: U3NeoSV1"
ORIGIN

Query Match 62.0%; Score 6.2; DB 9; Length 10;
Best Local Similarity 40.0%; Pred. No. 3.3e+06;
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 UGCDGGHNMD 10
              :||:||: ::
Db         10 TGCAGGACAT 1


RESULT 2
AJ594173
LOCUS       AJ594173                 13 bp    DNA     linear   GSS 15-JAN-2004
DEFINITION  Arabidopsis thaliana T-DNA flanking sequence, left border, clone
            393H08, genomic survey sequence.
ACCESSION   AJ594173
VERSION     AJ594173.1  GI:37943797
KEYWORDS    GSS; left border; T-DNA flanking sequence.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1
  AUTHORS   Brunaud,V., Balzergue,S., Dubreucq,B., Aubourg,S., Samson,F.,
            Chauvin,S., Bechtold,N., Cruaud,C., DeRose,R., Pelletier,G.,
            Lepiniec,L., Caboche,M. and Lecharny,A.
  TITLE     T-DNA integration into the Arabidopsis genome depends on sequences
            of pre-insertion sites
  JOURNAL   EMBO Rep. 3 (12), 1152-1157 (2002)
  MEDLINE   22363535
  PUBMED    12446565
REFERENCE   2  (bases 1 to 13)
  AUTHORS   Balzergue,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-OCT-2003) Balzergue S., UMRGV, INRA/CNRS, 2 rue
            Gaston Cremieux, 91057 Evry cedex, FRANCE
COMMENT     PCR was performed on DNA from transformants of Arabidopsis thaliana
            plants from INRA (Versailles). The DNA fragment(s) resulting from
            the PCR were directly sequenced from the left or the right border
            to determine the genomic sequence flanking the insertion. T-DNA
            derived sequences were removed. Information to order the
            corresponding mutant line and a link to a database providing a
            graphical display of the insertion site are available at
            http://dbsgap.versailles.inra.fr/publiclines/. This sequence has
            been generated in the framework of the French plant genomics
            program 'Genoplante' (http://www.genoplante.com and
            http://genoplante-info.infobiogen.fr).
FEATURES             Location/Qualifiers
     source          1. .13
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /cultivar="Wassillewskija"
                     /db_xref="taxon:3702"
                     /clone="393H08"
                     /clone_lib="Arabidopsis thaliana T-DNA insertion lines"
     misc_feature    1. .13
                     /note="T-DNA flanking sequence
                     left border"
ORIGIN

  Query Match             62.0%;  Score 6.2;  DB 9;  Length 13;
  Best Local Similarity   40.0%;  Pred. No. 3.4e+06;
  Matches    4;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 UGCDGGHNMD 10
              :||:||: ::
Db          1 TGCTGGTGAA 10


RESULT 3
AZ422762/c
LOCUS       AZ422762                 19 bp    DNA     linear   GSS 03-OCT-2000
DEFINITION  1M0201P12R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0201P12 R, genomic survey sequence.
ACCESSION   AZ422762
VERSION     AZ422762.1  GI:10546871
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 19)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,
            Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von
            Niederhausern,A. and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0201  row: P  column: 12
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 19.
FEATURES             Location/Qualifiers
     source          1. .19
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0201P12"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
ORIGIN

  Query Match             62.0%;  Score 6.2;  DB 8;  Length 19;
  Best Local Similarity   40.0%;  Pred. No. 3.5e+06;
  Matches    4;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 UGCDGGHNMD 10
              :||:||: ::
Db         18 TGCAGGCACA 9


RESULT 4
AZ791433
LOCUS       AZ791433                 19 bp    DNA     linear   GSS 16-FEB-2001
DEFINITION  2M0041A24F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC2M0041A24 F, genomic survey sequence.
ACCESSION   AZ791433
VERSION     AZ791433.1  GI:12934312
```

```
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 19)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,
            Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von
            Niederhausern,A. and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0041  row: A  column: 24
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 19.
FEATURES             Location/Qualifiers
     source          1. .19
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC2M0041A24"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
ORIGIN

  Query Match           62.0%;  Score 6.2;  DB 8;  Length 19;
  Best Local Similarity  40.0%;  Pred. No. 3.5e+06;
  Matches    4;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 UGCDGGHNMD 10
              :||:||: ::
Db          6 TGCAGGCACT 15


RESULT 5
AZ800056
LOCUS       AZ800056                 19 bp    DNA     linear   GSS 16-FEB-2001
DEFINITION  2M0057E20R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC2M0057E20 R, genomic survey sequence.
ACCESSION   AZ800056
VERSION     AZ800056.1  GI:12951797
```

```
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 19)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,
            Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von
            Niederhausern,A. and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0057  row: E  column: 20
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 19.
FEATURES             Location/Qualifiers
     source          1. .19
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC2M0057E20"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
ORIGIN

  Query Match           62.0%;  Score 6.2;  DB 8;  Length 19;
  Best Local Similarity  40.0%;  Pred. No. 3.5e+06;
  Matches    4;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 UGCDGGHNMD 10
              :||:||: ::
Db          3 TGCTGGAAAT 12


RESULT 6
CL435787
LOCUS       CL435787                 19 bp    DNA     linear   GSS 18-MAR-2004
DEFINITION  PST1409-NL.seq MICB1 Mus musculus genomic clone PST1409-NL.seq
            similar to Spats1, genomic survey sequence.
ACCESSION   CL435787
VERSION     CL435787.1  GI:45569831
```

```
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 19)
  AUTHORS   Hicks,G.G.
  TITLE     www.EScells.ca
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Hicks GG
            Mammalian Functional Genomics Centre
            Manitoba Institute of Cell Biology, Universitry of Manitoba
            ON5029, 675 McDermot Ave, Winnipeg, MB R3E 0V9, Canada
            Tel: 204 787 2133
            Fax: 204 787 2190
            Email: hicksgg@cc.umanitoba.ca
            U3NeoSV1 gene trap. Tag generated by plasmid rescue. Additional
            sequence information and target gene cloning can be generated. ES
            cell line harboring insertion mutation of target gene is available.
            Sequence analysis available from
            http://140.193.242.7/esdb/public_search_frame.php?PST=PST1409-NL.se
            q
            Class: Gene Trap.
FEATURES             Location/Qualifiers
     source          1. .19
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="129 sv"
                     /db_xref="taxon:10090"
                     /clone="PST1409-NL.seq"
                     /sex="Male"
                     /cell_type="Embryonic stem cell"
                     /cell_line="D3H (J1 subclone)"
                     /clone_lib="MICB1"
                     /note="Vector: U3NeoSV1"
ORIGIN

  Query Match             62.0%;  Score 6.2;  DB 9;  Length 19;
  Best Local Similarity   40.0%;  Pred. No. 3.5e+06;
  Matches    4;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 UGCDGGHNMD 10
              :||:||: ::
Db          4 TGCTGGTGAT 13


RESULT 7
AZ308311
LOCUS       AZ308311                 20 bp    DNA     linear   GSS 29-SEP-2000
DEFINITION  1M0011J12F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0011J12 F, genomic survey sequence.
ACCESSION   AZ308311
VERSION     AZ308311.1  GI:10348177
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 20)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,
            Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von
            Niederhausern,A. and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0011  row: J  column: 12
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 20.
FEATURES             Location/Qualifiers
     source          1. .20
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0011J12"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
ORIGIN

  Query Match             62.0%;  Score 6.2;  DB 8;  Length 20;
  Best Local Similarity   40.0%;  Pred. No. 3.5e+06;
  Matches    4;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 UGCDGGHNMD 10
              :||:||: ::
Db         11 TGCAGGCCAG 20


RESULT 8
AZ590589/c
LOCUS       AZ590589                 20 bp    DNA     linear   GSS 13-DEC-2000
DEFINITION  1M0400G15F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0400G15 F, genomic survey sequence.
ACCESSION   AZ590589
VERSION     AZ590589.1  GI:11712779
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 20)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,
            Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von
            Niederhausern,A. and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
```

```
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0400  row: G  column: 15
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 20.
FEATURES              Location/Qualifiers
     source           1. .20
                      /organism="Mus musculus"
                      /mol_type="genomic DNA"
                      /strain="C57BL/6J"
                      /db_xref="taxon:10090"
                      /clone="UUGC1M0400G15"
                      /sex="Male"
                      /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                      /clone_lib="Mouse 10kb plasmid UUGC1M library"
                      /note="Vector: PWD42nv; Purified genomic DNA from M.
                      musculus C57BL/6J (male) was obtained  from the Jackson
                      Laboratory Mouse DNA Resource
                      (http://www.jax.org/resources/documents/dnares/). The DNA
                      was hydrodynamically sheared by repeated passage through a
                      0.005 inch orifice at constant velocity. The sheared DNA
                      was blunt end-repaired with T4 DNA polymerase and T4
                      polynucleotide kinase. Adaptor oligonucleotides were
                      ligated to the blunt ends in high  molar excess. The
                      adaptored DNA was purified and size-selected for a 9.5 to
                      10.5 kb range using preparative agarose gel
                      electrophoresis. Vector DNA was prepared from a derivative
                      of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                      inducible derivative of plasmid R1. The vector was ligated
                      with adaptors complementary to the insert adaptors and
                      purified. The sheared, adaptored mouse DNA was annealed to
                      adaptored vector DNA, and transformed into
                      chemically-competent E. coli XL10-Gold (Stratagene) cells
                      and selected for ampicillin resistance."
ORIGIN

  Query Match             62.0%;  Score 6.2;  DB 8;  Length 20;
  Best Local Similarity   40.0%;  Pred. No. 3.5e+06;
  Matches    4;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 UGCDGGHNMD 10
              :||:||: ::
Db         20 TGCAGGATCA 11


RESULT 9
AZ630221
LOCUS       AZ630221                 20 bp    DNA     linear   GSS 13-DEC-2000
DEFINITION  1M0483K12R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0483K12 R, genomic survey sequence.
ACCESSION   AZ630221
VERSION     AZ630221.1  GI:11752411
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 20)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,
            Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von
            Niederhausern,A. and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0483  row: K  column: 12
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 20.
FEATURES              Location/Qualifiers
     source           1. .20
                      /organism="Mus musculus"
                      /mol_type="genomic DNA"
                      /strain="C57BL/6J"
                      /db_xref="taxon:10090"
                      /clone="UUGC1M0483K12"
                      /sex="Male"
                      /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                      /clone_lib="Mouse 10kb plasmid UUGC1M library"
                      /note="Vector: PWD42nv; Purified genomic DNA from M.
                      musculus C57BL/6J (male) was obtained  from the Jackson
                      Laboratory Mouse DNA Resource
                      (http://www.jax.org/resources/documents/dnares/). The DNA
                      was hydrodynamically sheared by repeated passage through a
                      0.005 inch orifice at constant velocity. The sheared DNA
                      was blunt end-repaired with T4 DNA polymerase and T4
                      polynucleotide kinase. Adaptor oligonucleotides were
                      ligated to the blunt ends in high  molar excess. The
                      adaptored DNA was purified and size-selected for a 9.5 to
                      10.5 kb range using preparative agarose gel
                      electrophoresis. Vector DNA was prepared from a derivative
                      of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                      inducible derivative of plasmid R1. The vector was ligated
                      with adaptors complementary to the insert adaptors and
                      purified. The sheared, adaptored mouse DNA was annealed to
                      adaptored vector DNA, and transformed into
                      chemically-competent E. coli XL10-Gold (Stratagene) cells
                      and selected for ampicillin resistance."
ORIGIN

  Query Match             62.0%;  Score 6.2;  DB 8;  Length 20;
  Best Local Similarity   40.0%;  Pred. No. 3.5e+06;
  Matches    4;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 UGCDGGHNMD 10
              :||:||: ::
Db          2 TGCTGGTGCA 11


RESULT 10
AZ789903
LOCUS       AZ789903                 20 bp    DNA     linear   GSS 16-FEB-2001
DEFINITION  2M0038F15F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC2M0038F15 F, genomic survey sequence.
ACCESSION   AZ789903
VERSION     AZ789903.1  GI:12931404
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 20)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,
            Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von
            Niederhausern,A. and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
```

Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0038 row: F column: 15
Seq primer: CGTTGTAAAACGACGGCCAGT
Class: plasmid ends
High quality sequence stop: 20.

FEATURES Location/Qualifiers
source 1. .20
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0038F15"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptored DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

ORIGIN

Query Match 62.0%; Score 6.2; DB 8; Length 20;
Best Local Similarity 40.0%; Pred. No. 3.5e+06;
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 1 UGCDGGHNMD 10
:||:||: ::
Db 8 TGCAGGCACT 17

RESULT 11
CO790328/c
LOCUS CO790328 21 bp mRNA linear EST 05-AUG-2004
DEFINITION NT009B_E10 St18-22 Neural tube (NT) Ambystoma mexicanum cDNA 5' similar to hypothetical protein, mRNA sequence.
ACCESSION CO790328
VERSION CO790328.1 GI:51006299
KEYWORDS EST.
SOURCE Ambystoma mexicanum (axolotl)
ORGANISM Ambystoma mexicanum
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Caudata; Salamandroidea; Ambystomatidae; Ambystoma.
REFERENCE 1 (bases 1 to 21)
AUTHORS Habermann,B., Bebin,A.G., Herklotz,S., Volkmer,M., Eckelt,K., Pehlke,K., Epperlein,H.H., Schackert,H.K., Wiebe,G. and Tanaka,E.M.
TITLE An Ambystoma mexicanum EST sequencing project: Analysis of 17,352 expressed sequence tags from embryonic and regenerating blastema cDNA libraries
JOURNAL Genome Biol. (2004) In press
COMMENT Contact: Elly M. Tanaka
Tanaka Lab
Max Planck Institute of Molecular Cell Biology and Genetics, Dresden
Pfotenhauerstrasse 108,01307 Dresden, Germany
Tel: 0049 351 210 2620
Fax: 0049 351 210 1489
Email: tanaka@mpi-cbg.de
Plate: NT009B row: 10 column: E
Seq primer: GCA CAT TAG GCC TAT TTA GGT GAC A.

FEATURES Location/Qualifiers
source 1. .21
/organism="Ambystoma mexicanum"
/mol_type="mRNA"
/db_xref="taxon:8296"
/tissue_type="Neural Tube, Notochord, Somites"
/cell_type="Includes Neural tube, notochord, somites"
/dev_stage="Stage 18-22"
/clone_lib="St18-22 Neural tube (NT)"
/note="Vector: pCMVSport6; Site_1: NotI; Site_2: SalI; Unnormalized cDNA plasmid library prepared by Invitrogen. Size fractionated mRNA was polydT primed and cloned into NotI-SalI site of pCMVSport6. Bacterial host is EMDH10B-TONA. Average insert size is 1.5 kB.
TAG_LIB=NT"

ORIGIN

Query Match 62.0%; Score 6.2; DB 7; Length 21;
Best Local Similarity 40.0%; Pred. No. 3.5e+06;
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 1 UGCDGGHNMD 10
:||:||: ::
Db 15 TGCAGGTGAA 6

RESULT 12
AZ594960
LOCUS AZ594960 21 bp DNA linear GSS 13-DEC-2000
DEFINITION 1M0407E16F Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC1M0407E16 F, genomic survey sequence.
ACCESSION AZ594960
VERSION AZ594960.1 GI:11717150
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 21)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0407 row: E column: 16
Seq primer: CGTTGTAAAACGACGGCCAGT
Class: plasmid ends
High quality sequence stop: 21.

FEATURES Location/Qualifiers
source 1. .21
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0407E16"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/note="Vector: PWD42nv; Purified genomic DNA from M.

musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptored DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

ORIGIN

Query Match 62.0%; Score 6.2; DB 8; Length 21;
Best Local Similarity 40.0%; Pred. No. 3.5e+06;
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 UGCDGGHNMD 10
              :||:||: ::
Db         12 TGCAGGCTAA 21
```

RESULT 13
AZ776814
LOCUS AZ776814 21 bp DNA linear GSS 16-FEB-2001
DEFINITION 2M0010I17R Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC2M0010I17 R, genomic survey sequence.
ACCESSION AZ776814
VERSION AZ776814.1 GI:12904767
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 21)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0010 row: I column: 17
Seq primer: CACACAGGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 21.
FEATURES Location/Qualifiers
source 1. .21
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0010I17"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/note="Vector: PWD42nv; Purified genomic DNA from M.

musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptored DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

ORIGIN

Query Match 62.0%; Score 6.2; DB 8; Length 21;
Best Local Similarity 40.0%; Pred. No. 3.5e+06;
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 UGCDGGHNMD 10
              :||:||: ::
Db         10 TGCTGGTACT 19
```

RESULT 14
AZ819369
LOCUS AZ819369 21 bp DNA linear GSS 20-FEB-2001
DEFINITION 2M0089O21R Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC2M0089O21 R, genomic survey sequence.
ACCESSION AZ819369
VERSION AZ819369.1 GI:12989277
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 21)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0089 row: O column: 21
Seq primer: CACACAGGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 21.
FEATURES Location/Qualifiers
source 1. .21
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0089O21"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/note="Vector: PWD42nv; Purified genomic DNA from M.

musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptored DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

ORIGIN

Query Match 62.0%; Score 6.2; DB 8; Length 21;
Best Local Similarity 40.0%; Pred. No. 3.5e+06;
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 UGCDGGHNMD 10
              :||:||: ::
Db          4 TGCTGGAACA 13
```

RESULT 15
AZ307716/c
LOCUS AZ307716 22 bp DNA linear GSS 29-SEP-2000
DEFINITION 1M0009N02R Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC1M0009N02 R, genomic survey sequence.
ACCESSION AZ307716
VERSION AZ307716.1 GI:10346985
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 22)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0009 row: N column: 02
Seq primer: CACACAGGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 22.
FEATURES Location/Qualifiers
source 1. .22
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0009N02"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/note="Vector: PWD42nv; Purified genomic DNA from M.

musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptored DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

ORIGIN

Query Match 62.0%; Score 6.2; DB 8; Length 22;
Best Local Similarity 40.0%; Pred. No. 3.5e+06;
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 UGCDGGHNMD 10
              :||:||: ::
Db         17 TGCAGGAGAG 8
```

Search completed: October 19, 2004, 16:07:31
Job time : 2183 secs